



UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO

---

FACULTAD DE CIENCIAS

PROCESOS OGARCH Y APLICACIONES A DATOS  
FINANCIEROS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:  
ACTUARIO

PRESENTA:  
PABLO ISAAC FRAGA ESPARZA

DIRECTOR DE TESIS:  
DRA. MA. ASUNCIÓN BEGOÑA FERNÁNDEZ  
FERNÁNDEZ



CIUDAD UNIVERSITARIA

2014



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# Hoja de Datos del Jurado

## 1. Datos del alumno

Apellido paterno  
Apellido materno  
Nombre(s)  
Teléfono  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Carrera  
Número de cuenta

## 2. Datos del tutor

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

## 3. Datos del sinodal 1

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

## 4. Datos del sinodal 2

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

## 5. Datos del sinodal 3

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

## 6. Datos del sinodal 4

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

## 7. Datos del trabajo escrito

Título  
Número de páginas  
Año

## 1. Datos del alumno

Fraga  
Esparza  
Pablo Isaac  
55 53 48 09  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Actuaría  
305553713

## 2. Datos del tutor

Dra.  
María Asunción Begoña  
Fernández  
Fernández

## 3. Datos del sinodal 1

Dra. en E.  
Ruth Selene  
Fuentes  
García

## 4. Datos del sinodal 2

Mat.  
Margarita Elvira  
Chávez  
Cano

## 5. Datos del sinodal 3

Dr.  
Fernando  
Baltazar  
Larios

## 6. Datos del sinodal 4

M. en C.  
Arrigo  
Coen  
Coria

## 7. Datos del trabajo escrito

Procesos OGARCH y aplicaciones a datos financieros  
100 p  
2014

“Was vorstellbar ist, ist auch machbar”  
Albert Einstein



# Agradecimientos

A Dios, porque sin él nada de esto sería posible.

A mi mamá y mi familia por su apoyo incondicional y confianza total a través de los años, estos han sido la base para lograr todos mis sueños. A Dani y a mis amigos de la Rondalla, Tone, Edgar, Santiago, Manuel, Zavala, Pedro, Warrior, Omaña, Taz y Pastor; por acompañarme en momentos difíciles e increíbles, por mantenerme en el camino cuando el futuro se veía incierto y por enseñarme que para lograr tus sueños, sólo necesitas tener fe.

A la UNAM y en especial a la Facultad de Ciencias, por abrirme sus puertas, ya que es un orgullo formar parte de esta institución, a la cual debo mi crecimiento académico y personal. A Jefferson, Marco e Irene, a través de sus clases me enseñaron a amar lo que hago. A Úrsula Iturrarán, Patricia Luna, Jorge y Jorge, Nils Ackermann, Guadalupe Lucio y Leticia Cid, por su confianza y por el apoyo que me brindaron durante mi estadia en la UNAM.

De manera muy especial a mi asesora Begoña Fernández por la paciencia que mostró durante varios meses y por ayudarme en la realización de esta tesis. A mis sinodales Ruth Fuentes, Margarita Chávez, Fernando Baltazar y Arrigo Coen por su tiempo y su dedicación en la revisión de esta tesis; sus comentarios y aportaciones hicieron que fuera un trabajo más consistente. A Arrigo agradezco también por sus clases y su paciencia durante estos últimos años. A Nelson Muriel, por las charlas y discusiones que tuvimos, éstas ayudaron de manera sustancial para la comprensión de los temas de esta tesis.

A Pilar. Gracias por tu amistad, apoyo, compañía y por las invaluable lecciones que me enseñaste, me ayudaste a convertirme en un mejor hombre y mejor ser humano.

# Índice general

<b>Agradecimientos</b>	<b>v</b>
<b>Introducción</b>	<b>ix</b>
<b>1. Preliminares</b>	<b>1</b>
1.1. Componentes principales de la muestra . . . . .	5
1.2. Componentes principales de variables estandarizadas . . . . .	10
<b>2. Series de tiempo ARCH y GARCH</b>	<b>17</b>
2.1. Fundamentos del análisis de las series de tiempo . . . . .	18
2.2. Modelos de series de tiempo ARCH y GARCH . . . . .	21
2.2.1. Procesos ARCH . . . . .	21
2.2.2. El modelo ARCH(1) . . . . .	22
2.2.3. Procesos GARCH . . . . .	23
2.2.4. Ajuste de modelos GARCH . . . . .	25
2.3. Media Móvil con Ponderación Exponencial . . . . .	27
2.4. Fundamentos de las series de tiempo multivariadas . . . . .	28
2.5. Procesos GARCH Multivariados . . . . .	31
2.5.1. Estructura general de los modelos . . . . .	32
2.5.2. Modelos de correlación condicional . . . . .	34
2.5.3. Modelos de covarianza condicional . . . . .	37
2.5.4. Ajuste de modelos GARCH Multivariados . . . . .	40
2.5.5. PC-GARCH y O-GARCH . . . . .	41
<b>3. Aplicación del modelo O-GARCH</b>	<b>45</b>
3.1. <i>Forwards</i> sobre el <i>West Texas Intermediate</i> . . . . .	45
3.1.1. Aplicación . . . . .	46
3.2. <i>Forwards</i> sobre el <i>Brent Blend</i> . . . . .	64

3.2.1. Aplicación . . . . .	64
<b>4. Conclusiones</b>	<b>73</b>
<b>A. Teoremas</b>	<b>75</b>
<b>B. Códigos</b>	<b>79</b>

# Introducción

Modelar la volatilidad ha sido objeto de mucha atención desde que Engle propuso los modelos ARCH en 1982 en su trabajo *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*. A raíz de esto, se propusieron diversas extensiones y variantes del modelo, siendo algunas de ellas los modelos GARCH, que fueron propuestos por Bollerslev en 1986 en su trabajo *Generalized autoregressive conditional heteroskedasticity*, y los modelos GARCH Multivariados.

Si bien, modelar la volatilidad de los rendimientos ha sido de mucha atención durante los últimos años, era de esperarse que también se haya vuelto sumamente importante la comprensión de los movimientos conjuntos de estos, es por ello que se ha prestado tanta atención a los modelos GARCH Multivariados.

Este trabajo abarca, principalmente, un modelo GARCH Multivariado, llamado GARCH Ortogonal (O-GARCH), que fue propuesto por Alexander C. y Chibumba A. en 1996 en su trabajo *Multivariate Orthogonal Factor GARCH*. Este modelo hace uso del análisis de componentes principales, pero a diferencia del GARCH de Componentes Principales (PC-GARCH), se reduce la dimensión del problema al usar sólo las componentes principales que explican la mayor parte de la variabilidad del sistema original.

El modelo O-GARCH, a diferencia de algunos otros modelos GARCH Multivariados, hace uso de los modelos GARCH univariados para estimar la matriz de covarianzas.

Los modelos GARCH Multivariados, presentan 2 grandes desventajas. La primera es el número de parámetros, pues conforme aumenta la dimensión,

aumentarán los parámetros llegando a ser en algunos casos un número excesivamente grande.

El otro problema que presentan es al momento del ajuste. Porque nos gustaría hacer uso de la función de verosimilitud, como pasa en el caso univariado, pero para el caso multivariado se vuelve cada vez más complicado conforme aumenta la dimensión del modelo. Es por ello que se tiene que recurrir a otra forma de ajuste, y ésta es el ajuste por etapas. En este tipo de ajuste, como veremos más adelante en los modelos GARCH Multivariados, se va revisando que el ajuste sea correcto en diferentes partes del modelo.

La ventaja del O-GARCH, es que no tiene este tipo de problemas, ya que al hacer uso del análisis de componentes principales, estamos reduciendo la dimensión de los datos; y cuenta con un ajuste por etapas relativamente sencillo, puesto que se trabaja con modelos GARCH(1,1).

En el Capítulo 1 presentamos de manera detallada el análisis de componentes principales, su implementación y su ajuste a los datos. En el Capítulo 2 abarcamos las propiedades y características de los modelos ARCH y GARCH univariados, prestando especial atención a los modelos de primer orden, es decir al ARCH(1) y al GARCH(1,1). Posteriormente, analizamos algunos de los más importantes y más utilizados modelos GARCH Multivariados, y ahí mismo vemos las desventajas que presentan estos modelos cuando queremos ajustarlos a datos de la vida real. Para finalizar, aplicamos este modelo a los precios de los *forwards* sobre el petróleo crudo WIT y a los precios de los *forwards* sobre el petróleo crudo *Brent Blend*.

# Capítulo 1

## Preliminares

El contenido de este capítulo está basado en Jolliffe (2002); Johnson y Wichern (2007); Härdle y Simar (2003); y McNeil, Frey y Embrechts (2005).

El objetivo del análisis de componentes principales (ACP), es reducir la dimensión de un conjunto de datos altamente correlacionados entre sí, mediante la búsqueda de un pequeño número de combinaciones lineales, no correlacionadas entre sí, que conserven la mayor cantidad de varianza de los datos originales. Esto se logra haciendo una transformación a un nuevo conjunto de variables, las componentes principales, que no están correlacionadas entre sí y se encuentran ordenadas de modo que la primeras componentes tengan la mayor parte de la varianza de las variables originales.

Para hacer esto, supongamos que  $X = (X_1, \dots, X_p)'$  es un vector de  $p$  variables aleatorias y que el vector de medias y la matriz de covarianzas de  $X$  existen. Deseamos encontrar una combinación lineal

$$\delta'X = \delta_1 X_1 + \dots + \delta_p X_p, \quad (1.1)$$

de las componentes de  $X$  tal que la varianza de (1.1) sea lo más grande posible. Dado que esta varianza puede ser incrementada multiplicando el vector  $\delta$  por una constante, es usual restringir a que  $\delta$  sea un vector unitario, es decir,  $\delta'\delta = 1$ , o bien que

$$\sum_{j=1}^p \delta_j^2 = 1.$$

Cabe resaltar que ésta es la proyección del vector de variables aleatorias  $X$  sobre la recta  $\delta$ . Ahora, si sacamos la varianza de (1.1) y después maximizamos, obtenemos

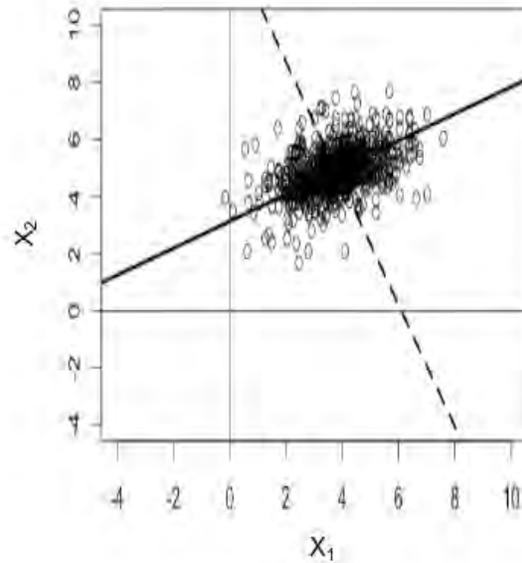
$$\max_{\{\delta: \|\delta\|=1\}} \text{Var}(\delta'X) = \max_{\{\delta: \|\delta\|=1\}} \delta' \text{Var}(X) \delta, \quad (1.2)$$

usando el Teorema A.4, se maximiza cuando la dirección  $\delta$  está dada por el eigenvector  $\gamma_1$  que corresponde al eigenvalor más grande  $\lambda_1$  de la matriz de covarianzas  $\Sigma = \text{Var}(X)$ .

Si seguimos este mismo procedimiento y buscamos la dirección ortogonal al primer eigenvector que maximiza la varianza, nos dará el eigenvector correspondiente al segundo eigenvalor más grande  $\lambda_2$  de la matriz de covarianzas, y así sucesivamente hasta encontrar las  $p$  direcciones, que serán ortogonales entre sí, y que coincidirán con los eigenvectores correspondientes a cada eigenvalor de la matriz de covarianzas. Estas combinaciones lineales son las componentes principales y tienen la siguiente expresión  $y_k = \gamma_k'X$  con  $k = 1, \dots, p$ .

Una vez que sabemos cómo encontrar analíticamente las componentes principales, observemoslo de manera gráfica en  $\mathbb{R}^2$ . En la Figura 1.1 tenemos observaciones de una distribución normal bivariada, con centro en  $(3.7845879, 4.907856)$  y dos líneas, una continua y la otra punteada. La línea continua representa el vector en el cual encontramos la varianza máxima al hacer la proyección. Éste es el eigenvector correspondiente al eigenvalor más grande de la matriz de covarianzas de las observaciones; mientras que la punteada es el vector que maximiza la varianza en la dirección ortogonal al primer eigenvector, cuando hacemos la proyección sobre él. Este vector es el eigenvector correspondiente al segundo eigenvalor más grande de la matriz de covarianzas de las observaciones.

Ello sugiere la interpretación del análisis de componentes principales como un cambio de base, o como una técnica de rotación de datos. Por lo general se suelen centrar las observaciones antes del análisis, para que el origen de ambos sistemas coincida. En la Figura 1.2 se puede apreciar las mismas observaciones pero vistas desde el sistema de coordenadas de nuestras com-



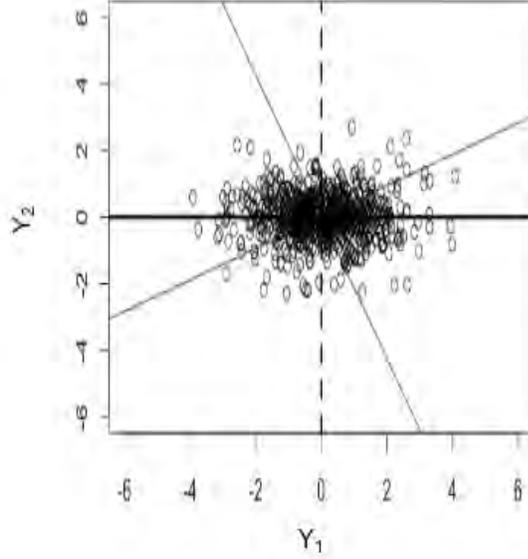
**Figura 1.1:** Observaciones con media en  $(3.7845879, 4.907856)$ .

ponentes principales.

En términos matriciales, si tenemos el vector aleatorio  $X$  con vector de medias  $\mu$  y matriz de covarianzas  $\Sigma$  con su descomposición espectral  $\Sigma = \Gamma\Lambda\Gamma'$ , donde  $\Lambda$  es la matriz diagonal de los eigenvalores y  $\Gamma$  es la matriz de los eigenvectores; la matriz de componentes principales  $Y$  estará definida como

$$Y := \Gamma'(X - \mu). \quad (1.3)$$

**Observación 1.1.** *A la matriz que contiene los eigenvectores se le conoce en la literatura como la matriz de “Loadings” o “Coefficients”; y a la matriz de las componentes principales se lo conoce como la matriz de “Scores”.*



**Figura 1.2:** Observaciones con media en (0,0) del sistema original.

**Observación 1.2.** La matriz  $Y$  antes mencionada cumplirá que

$$\begin{aligned}
 \mathbb{E}[Y_i] &= 0, & \forall i = 1, \dots, p. \\
 \text{Var}(Y_i) &= \lambda_i, & \forall i = 1, \dots, p. \\
 \text{Cov}(Y_i, Y_j) &= 0, & i \neq j. \\
 \text{Var}(Y_1) &\geq \dots \geq \text{Var}(Y_p) \geq 0. \\
 \sum_{i=1}^p \text{Var}(Y_i) &= \text{tr}(\Sigma). \\
 \prod_{i=1}^p \text{Var}(Y_i) &= |\Sigma|.
 \end{aligned}$$

Entonces la proporción de la varianza que es explicada por la  $k$ -ésima componente es

$$\frac{\lambda_k}{\lambda_1 + \dots + \lambda_p}, \quad \forall k = 1, \dots, p.$$

La expresión (1.3) es invertible, por lo que  $X = \mu + \Gamma Y$ , donde  $Y$  tiene las componentes principales ordenadas de mayor a menor, de acuerdo a su varianza.

Supongamos ahora que creemos que las primeras  $k$  componentes explican la mayor parte de la variabilidad de  $X$ . Podemos ver a  $Y$  como  $(Y_1', Y_2')'$ , donde  $Y_1 \in \mathbb{R}^k$  y  $Y_2 \in \mathbb{R}^{p-k}$ ; y a  $\Gamma$  como  $(\Gamma_1, \Gamma_2)$  donde  $\Gamma_1 \in \mathbb{R}^{p \times k}$  y  $\Gamma_2 \in \mathbb{R}^{p \times (p-k)}$ . Obteniendo como resultado

$$X = \mu + \Gamma_1 Y_1 + \Gamma_2 Y_2 = \mu + \Gamma_1 Y_1 + \varepsilon, \quad (1.4)$$

donde  $\Gamma_2 Y_2$  se puede pensar como un error; ya que su matriz de covarianzas tiene entradas muy pequeñas en comparación con la matriz de covarianzas de  $\Gamma_1 Y_1$ .

## 1.1. Componentes principales de la muestra

Supongamos que tenemos observaciones de nuestros datos multivariados  $X_1, \dots, X_p$  idénticamente distribuidas y con matriz de covarianzas  $\Sigma$  desconocida. Estimamos a  $\Sigma$  con la matriz de covarianzas muestral  $S$ , donde

$$S := \frac{1}{p} \sum_{i=1}^p (X_i - \bar{X})(X_i - \bar{X})', \quad y \quad \bar{X} := \frac{1}{p} \sum_{i=1}^p X_i.$$

La descomposición espectral de la matriz  $S_x$ , que es simétrica y positiva semidefinida, resulta

$$S_x = GLG',$$

donde  $G$  es la matriz de eigenvectores de  $S_x$  y  $L = \text{diag}(l_1, \dots, l_p)$  es la matriz diagonal de los eigenvalores ordenados. Así la matriz de las componentes principales nos queda  $Y = G'(X - \bar{X})$  y cumple que su matriz de covarianzas muestral es  $L$ , como se presenta a continuación

$$\begin{aligned} S_y &= \frac{1}{p} \sum_{i=1}^p (Y_i - \bar{Y})(Y_i - \bar{Y})' = \frac{1}{p} \sum_{i=1}^p Y_i Y_i' \\ &= \frac{1}{p} \sum_{i=1}^p G'(X_i - \bar{X})(X_i - \bar{X})'G = G' S_x G = L. \end{aligned}$$

Ilustrando lo anterior, considérense los log-rendimientos de 10 acciones del CAC 40: Air Liquide, Paribas, Pernord Ricard, L'Oréal, SocGen, Total, Michelin, Danone, CapGemini y BIC; los datos fueron recabados diariamente del 01/01/1994 al 09/02/1999 y tomados de la página personal de la Dra. Carol Alexander <http://www.carolalexander.org/discussionpapers.php>. Quien trabaja con estos datos en Alexander (2001) pero no realiza el siguiente análisis.

	alq	pbs	peu	orl	soc	tot	mic	dan	cap	bic
alq	1									
pbs	0.3894	1								
peu	0.4204	0.4062	1							
orl	0.5097	0.4721	0.4285	1						
soc	0.4227	0.6150	0.4546	0.4744	1					
tot	0.4452	0.3741	0.3529	0.4192	0.3437	1				
mic	0.3518	0.3542	0.4136	0.4350	0.3949	0.3349	1			
dan	0.4167	0.3952	0.3622	0.4914	0.3908	0.3963	0.3559	1		
cap	0.2375	0.2740	0.3497	0.3006	0.3077	0.1896	0.2820	0.2211	1	
bic	0.3104	0.2491	0.2420	0.3640	0.2749	0.2694	0.2759	0.3065	0.1963	1

**Tabla 1.1:** Matriz de correlaciones de los log-rendimientos de las acciones analizadas.

La Tabla 1.1 muestra la falta de una correlación fuerte entre los datos; en cambio si observamos la Tabla 1.2, las primeras 6 componentes principales explican una varianza del 83.22%. Esto nos lleva a cuestionar la certeza de que con 6 componentes principales es suficiente.

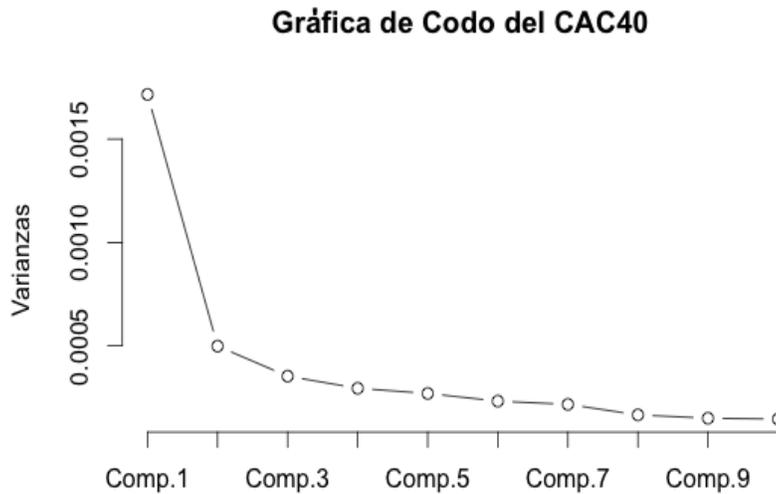
Entonces, ¿cuál es el número de componentes principales que deben usarse? Existe un método visual bastante útil para determinar el número apropiado de componentes principales, el cual consiste en buscar una especie de “codo” en la gráfica (Gráfica de Codo) y que el resto de los eigenvalores sean de la misma magnitud.

En la Figura 1.3, podemos notar que a partir del quinto o sexto eigenvalor el segmento de recta es casi horizontal y el resto de los eigenvalores tienen una magnitud muy parecida.

Ahora, veamos la base de datos de los vencimientos mensuales de los *forwards* del petróleo crudo del WTI entre 1 mes y 12 meses; la cual fue recabada diariamente del 04/02/1993 al 24/03/1999. Ésta también fue tomada

$\lambda_i$	Proporción Varianza	Proporción Acumulada
0.0017160469	0.42451980	0.42451980
0.0004985639	0.12333592	0.54785572
0.0003531877	0.08737241	0.63522813
0.0002945114	0.07285694	0.70808507
0.0002695146	0.06667316	0.77475823
0.0002325059	0.05751787	0.83227610
0.0002166951	0.05360656	0.88588266
0.0001658784	0.04103540	0.92691806
0.0001497184	0.03703770	0.96395575
0.0001457026	0.03604425	1.00000000

**Tabla 1.2:** Eigenvalores de las componentes principales de los log-rendimientos de las acciones analizadas.



**Figura 1.3:** Gráfica de Codo de los log-rendimientos de las acciones analizadas.

de la página personal de la Dra. Carol Alexander, quien trabaja con estos datos en Alexander (2001) pero no realiza el siguiente análisis.

En la Tabla 1.3 se puede ver que los datos están fuertemente correlacionados, lo que nos lleva a suponer un mejor desempeño del análisis de componentes principales.

	1m	2m	3m	4m	5m	6m	7m	8m	9m	10m	11m	12m
1m	1											
2m	0.9500	1										
3m	0.9295	0.9936	1									
4m	0.9105	0.9827	0.9960	1								
5m	0.8928	0.9702	0.9884	0.9974	1							
6m	0.8781	0.9585	0.9801	0.9925	0.9983	1						
7m	0.8646	0.9468	0.9709	0.9860	0.9944	0.9984	1					
8m	0.8517	0.9349	0.9609	0.9783	0.9889	0.9949	0.9986	1				
9m	0.8382	0.9225	0.9503	0.9698	0.9825	0.9903	0.9957	0.9988	1			
10m	0.8250	0.9103	0.9395	0.9608	0.9752	0.9846	0.9915	0.9961	0.9989	1		
11m	0.8083	0.8941	0.9248	0.9476	0.9637	0.9747	0.9832	0.9893	0.9934	0.9960	1	
12m	0.7975	0.8841	0.9160	0.9402	0.9578	0.9701	0.9797	0.9870	0.9923	0.9961	0.9960	1

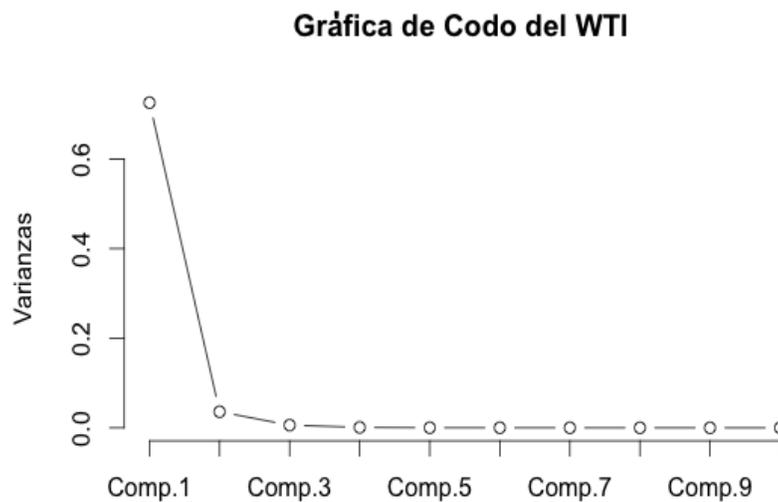
**Tabla 1.3:** Matriz de correlaciones de los log-rendimientos de los vencimientos mensuales de los *forwards* del WTI.

$\lambda_i$	Proporción Varianza	Proporción Acumulada
2.312412e-03	0.947045380	0.94704538
1.061673e-04	0.043480690	0.99052608
1.868485e-05	0.007652357	0.99817843
2.679683e-06	0.001097461	0.99927589
6.722541e-07	0.000275320	0.99955121
4.583032e-07	0.000187697	0.99973891
2.948113e-07	0.000120739	0.99985965
1.547259e-07	0.000063367	0.99992302
8.352090e-08	0.000034205	0.99995720
4.799800e-08	0.000019657	0.99997690
3.503763e-08	0.000014349	0.99999120
2.140318e-08	0.000008765	1

**Tabla 1.4:** Eigenvalores de las componentes principales de los log-rendimientos de los vencimientos mensuales de los *forwards* del WTI.

En la Tabla 1.4 observamos que la proporción de la varianza que explica la primer componente principal es cercana al 95% y con 2 componentes

principales podemos explicar el 99%. Es por ello que el análisis de componentes principales se vuelve más efectivo entre más correlacionadas estén las variables.



**Figura 1.4:** Gráfica de Codo de los log-rendimientos de los vencimientos mensuales de los *forwards* del WTI.

En la Figura 1.4 basta con usar las primeras 2 o inclusive 3 componentes principales, donde encontramos esa especie de “codo”, porque después de la tercer componente los datos ya son muy parecidos en magnitud.

En conclusión, el análisis de componentes principales se puede realizar a cualquier conjunto de datos; pero como ya se ejemplificó, entre menor correlación tengan, más componentes estaremos usando y menos útil se vuelve el método, por lo que es sumamente importante revisar la correlación de los datos antes de hacer el análisis.

## 1.2. Componentes principales de variables estandarizadas

El análisis de componentes principales es sensible a cambios de escala, lo cual significa que si multiplicamos una variable por un escalar, obtenemos diferentes eigenvalores y eigenvectores. Esto se debe a que los eigenvalores se obtienen de la matriz de covarianzas y no de la matriz de correlación. Supongamos que  $X = (X_1, \dots, X_p)'$  es un vector de  $p$  variables aleatorias y que  $\mu$  el vector de medias y  $\Sigma$  la matriz de covarianzas de  $X$  existen, entonces

$$\text{Var}(X - \mu) = \text{Var}(X) = \Sigma,$$

por lo que los eigenvectores resultantes de centrar las observaciones serán los mismos que los de  $\Sigma$ . No obstante, esto no sucede cuando multiplicamos por un escalar. Supongamos que  $A$  es una matriz diagonal con entradas reales  $a_1, \dots, a_p$ , entonces

$$\text{Var}(AX) = A\text{Var}(X)A' = A\Sigma A',$$

por lo que los eigenvectores de  $\text{Var}(AX)$  serán totalmente diferentes a los de  $\Sigma$ .

La dinámica es totalmente diferente con la matriz de correlaciones, porque para obtenerla, siempre nos fijamos en la matriz de covarianzas y hacemos el siguiente producto

$$\frac{\sigma_{i,j}}{\sigma_i \sigma_j},$$

obteniendo

$$\rho_{i,j}(\Sigma) = \frac{\sigma_{i,j}}{\sigma_i \sigma_j} = \frac{a_i a_j \sigma_{i,j}}{a_i \sigma_i a_j \sigma_j} = \rho_{i,j}(A\Sigma A'), \quad \forall i, j = 1, \dots, p,$$

donde  $\rho_{i,j}(\cdot)$  es el coeficiente de correlación entre la variable  $X_i$  y  $X_j$  calculado a partir de la matriz de covarianzas.

Así, vemos claramente que la matriz de correlaciones es invariante ante cambios de escala, y por ende el análisis de componentes principales sobre la matriz de correlaciones será el mismo con las observaciones a diferente escala.

## 1.2. COMPONENTES PRINCIPALES DE VARIABLES ESTANDARIZADAS 11

Ahora, si estandarizamos el vector  $X$ , tendremos que el vector  $Z$ ,

$$\begin{aligned} Z_1 &:= \frac{(X_1 - \mu_1^X)}{\sigma_1^X}, \\ &\vdots \\ Z_p &:= \frac{(X_p - \mu_p^X)}{\sigma_p^X}. \end{aligned}$$

O en notación matricial,

$$Z = (V_X^{1/2})^{-1}(X - \mu_X),$$

donde,

$$V_X^{1/2} = \begin{bmatrix} \sigma_1^X & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_p^X \end{bmatrix}.$$

Resulta obvio que  $\mathbb{E}[Z] = 0$ , en tanto que

$$\begin{aligned} \text{Var}(Z) &= \text{Var}\left((V_X^{1/2})^{-1}(X - \mu_X)\right) = (V_X^{1/2})^{-1} \Sigma_X (V_X^{1/2})^{-1} \\ &= \begin{bmatrix} 1 & \cdots & \frac{\sigma_1^X \sigma_p^X \sigma_{1,p}^X}{\sigma_1^X \sigma_1^X \sigma_p^X \sigma_p^X} \\ \vdots & \ddots & \vdots \\ \frac{\sigma_1^X \sigma_p^X \sigma_{1,p}^X}{\sigma_1^X \sigma_1^X \sigma_p^X \sigma_p^X} & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 & \cdots & \frac{\sigma_{1,p}^X}{\sigma_1^X \sigma_p^X} \\ \vdots & \ddots & \vdots \\ \frac{\sigma_{1,p}^X}{\sigma_1^X \sigma_p^X} & \cdots & 1 \end{bmatrix}, \end{aligned}$$

y por otro lado tenemos que

$$\rho_{i,j}^Z = \frac{\sigma_{i,j}^Z}{\sigma_i^Z \sigma_j^Z} = \frac{\sigma_{i,j}^X}{\sigma_i^X \sigma_j^X} = \rho_{i,j}^X, \quad \forall i, j = 1, \dots, p,$$

entonces,

$$\text{Var}(Z) = \rho(Z) = \rho(X). \quad (1.5)$$

Por eso, cuando queramos obtener las componentes principales de las variables estandarizadas  $Z$ , basta con que hagamos el análisis de componentes principales pero sobre la matriz de correlación de  $X$ .

**Observación 1.3.** *Los eigenvectores y eigenvalores obtenidos de la matriz  $\Sigma$  no son en general los mismos que se obtienen a partir de la matriz  $\rho$ .*

Las propiedades de la Observación 1.2 también se cumplirán para las componentes principales obtenidas de la matriz de correlación. En este caso,  $tr(\Sigma) = p$  por lo que la varianza total que explica cada componente es

$$\frac{\lambda_k}{p}, \quad \forall k = 1, \dots, p,$$

donde  $\lambda_k$  son los eigenvalores de la matriz de correlación  $\rho$ .

Al hacer el análisis de esta manera, existe un criterio llamado *Regla de Kaiser*, que dice que nos debemos quedar con aquellas componentes cuya varianza sea mayor a 1, es decir, con aquellas que expliquen más de  $1/p$  de la varianza total. Sin embargo, este método no resulta ser tan bueno como la Gráfica de Codo y no debe ser usado a la ligera.

En general, se estandariza para poder comparar variables que están en diferentes unidades.

**Observación 1.4.** *El análisis de componentes principales sólo debe ser aplicado a datos que tienen aproximadamente la misma escala en cada una de sus variables.*

**Ejemplo 1.1.** *Usaremos los datos recabados por el Banco Nacional Suizo, donde se midieron 200 billetes viejos de mil francos, de los cuales 100 eran genuinos y 100 eran falsos. Se tomaron 6 medidas por billete, siendo éstas nuestras 6 variables.*

- $X_1$  : Longitud del billete medido por el centro del mismo,
- $X_2$  : Altura del billete medida por la izquierda del mismo,
- $X_3$  : Altura del billete medida por la derecha del mismo,
- $X_4$  : Distancia desde el marco interior al borde inferior,
- $X_5$  : Distancia desde el marco interior al borde superior,
- $X_6$  : Longitud de la diagonal del billete.

*Estas medidas están dadas en milímetros.*

## 1.2. COMPONENTES PRINCIPALES DE VARIABLES ESTANDARIZADAS13

Tomamos estos datos de la base *ForgedBankNotes*, que se puede descargar de *R Project*. Aquí reproducimos algunos de los resultados que se pueden encontrar en Flury y Riedwyl (1988).

En este ejemplo no estandarizaremos los datos. El vector de medias es

$$\mu = (214.8960, 130.1215, 129.9565, 9.4175, 10.6505, 140.4835)'$$

el vector de eigenvalores es

$$\Lambda = (2.9853, 0.9309, 0.2421, 0.1936, 0.0847, 0.0353)'$$

y la matriz de eigenvectores es

$$\Gamma = \begin{bmatrix} -0.0437 & 0.0107 & -0.3263 & 0.5616 & 0.7525 & 0.0980 \\ 0.1121 & 0.0714 & -0.2589 & 0.4554 & -0.3468 & -0.7665 \\ 0.1391 & 0.0662 & -0.3447 & 0.4153 & -0.5346 & 0.6316 \\ 0.7683 & -0.5630 & -0.2180 & -0.1861 & 0.0999 & -0.0222 \\ 0.2017 & 0.6592 & -0.5566 & -0.4506 & 0.1019 & -0.0348 \\ -0.5789 & -0.4885 & -0.5917 & -0.2584 & -0.0844 & -0.0456 \end{bmatrix}.$$

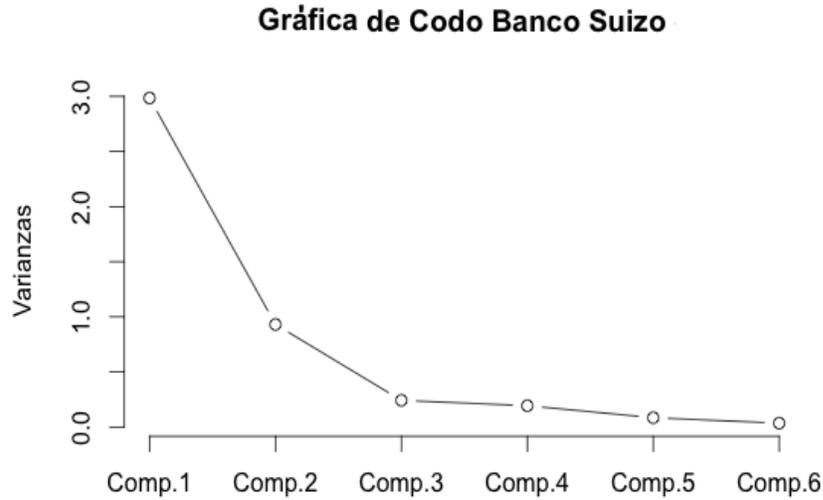
Realizamos el análisis de componentes principales y hacemos la Gráfica de Codo, la cual mostramos en la Figura 1.5. Al ver la Figura, no podríamos estar seguros de quedarnos con 3 componentes, porque las magnitudes de la tercer y de la cuarta componente son bastante parecidas, quizás lo mejor sea quedarse con 4 o hasta 5 componentes.

Veremos ahora lo sensible que se vuelve el análisis de componentes principales a un cambio de escala en las variables. Supongamos que las variables  $X_1$ ,  $X_2$ ,  $X_3$ , y  $X_6$  están medidas en milímetros y que  $X_4$  y  $X_5$  están en centímetros, al hacer el cambio de escala, nos queda que el vector de medias es

$$\mu = (21.4896, 13.0121, 12.9956, 9.4175, 10.6505, 14.0483)'$$

el vector de eigenvalores es

$$\Lambda = (2.1012, 0.6262, 0.0046, 0.0020, 0.0008, 0.0003)'$$



**Figura 1.5:** Gráfica de Codo de los datos del Banco Nacional Suizo.

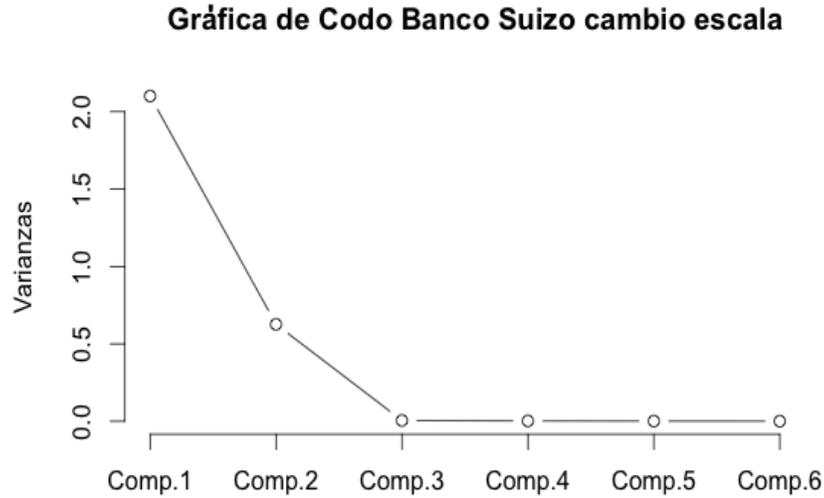
y la matriz de eigenvectores es

$$\Gamma = \begin{bmatrix} -0.0049 & 0.0010 & -0.0541 & 0.6324 & 0.7661 & -0.0999 \\ 0.0107 & -0.0127 & 0.1003 & 0.5351 & -0.3343 & 0.7690 \\ 0.0141 & -0.0154 & 0.0629 & 0.5569 & -0.5373 & -0.6298 \\ 0.9920 & 0.1172 & -0.0453 & -0.0062 & 0.0085 & 0.0020 \\ 0.1134 & -0.9905 & -0.0762 & -0.0117 & 0.0068 & 0.0030 \\ -0.0520 & 0.0687 & -0.9875 & 0.0563 & -0.1111 & 0.0431 \end{bmatrix}.$$

Comparando las columnas de la matriz de *Loadings* de antes y después del cambio de escala, podemos observar que ahora la primer componente estará dominada por  $X_4$ , la segunda componente estará dominada por  $X_5$  y la tercer componente estará dominada por  $X_6$  mientras que el resto de las variables, en los tres casos, tendrán un peso menor.

En cambio, la cuarta, quinta y sexta componente, permanecen casi iguales. Por lo cual, al hacer un cambio de escala en las observaciones, las componentes principales que explican la mayor parte de la varianza presenta un

## 1.2. COMPONENTES PRINCIPALES DE VARIABLES ESTANDARIZADAS15



**Figura 1.6:** Gráfica de Codo de los datos del Banco Suizo con cambio de escala.

cambio drástico.

En la Figura 1.6, podemos notar que basta con tomar 2 componentes principales para el análisis. Aquí es un buen momento para utilizar la Regla de Kaiser, en este caso sólo nos quedaríamos con la primer componente porque es la única cuya varianza es mayor a 1, pero esta componente sólo explica el 76.81 %, mientras que si tomamos las primeras dos componentes estamos explicando una varianza del 99.7 %. Podemos concluir que el método no es tan bueno como el Gráfica de Codo y siempre hay que tomar en cuenta otros factores para poder usarlo.



# Capítulo 2

## Series de tiempo ARCH y GARCH

El contenido de este capítulo está basado en McNeil, Frey y Embrechts (2005); Brockwell y Davis (2002); Zakoïan y Francq (2010); Alexander (2001); y Alexander (2002).

En este capítulo veremos los conceptos esenciales del análisis de las series de tiempo, tales como la estacionariedad, la autocorrelación y su estimación, procesos de ruido blanco y analizaremos los procesos univariados ARCH y GARCH para ajustar la volatilidad, porque se ha probado que los procesos GARCH son fáciles de ajustar y son la base para el entendimiento de modelos más complicados; como por ejemplo los modelos GARCH Multivariados.

Existen algunas propiedades resultantes de observaciones empíricas, que parecen ser comunes a la mayoría de las series de tiempo financieras. A estas propiedades las definiremos como *hechos estilizados* y las enunciamos a continuación.

- (a) La serie de los rendimientos no son i.i.d aunque muestran una autocorrelación pequeña.
- (b) Las series de rendimientos absolutos y cuadrados deben mostrar una fuerte autocorrelación.

- (c) La esperanza condicionada a la información pasada de los rendimientos debe ser cercana a cero.
- (d) La volatilidad de los rendimientos varía en el tiempo.
- (e) La distribución de los rendimientos es *leptokurtica* o de cola pesada.
- (f) Los valores extremos de los rendimientos aparecen en *clusters*.

**Observación 2.1.** *Cuando usamos el término clusters, nos referimos a la tendencia de que los rendimientos extremos sean seguidos por otros rendimientos extremos, aunque no con el mismo signo.*

Muchos análisis empíricos sugieren que la distribución de los rendimientos diarios, o de intervalos de tiempo aún más pequeños, tienen colas que decaen lentamente, en vez de exponencialmente como lo hacen las colas de una distribución normal. Además, cabe resaltar que conforme incrementemos el intervalo de los rendimientos, pasando de los diarios a los semanales, mensuales o anuales, el fenómeno que hasta el momento hemos identificado tiende a ser menos pronunciado; la volatilidad en los *clusters* disminuye y los rendimientos empiezan a verse i.i.d. y con colas menos pesadas.

## 2.1. Fundamentos del análisis de las series de tiempo

En esta sección abordaremos algunos resultados esenciales de las series de tiempo univariadas, haciendo énfasis en los puntos relevantes para modelar los rendimientos.

Un modelo de serie de tiempo es un proceso estocástico  $(X_t)_{t \in \mathbb{Z}}$ , es decir, una familia de variables aleatorias indexada en los enteros y definida en algún espacio de probabilidad  $(\Omega, \mathcal{F}, \mathcal{P})$ .

Asumiendo que existen los primeros 2 momentos de una serie de tiempo, definimos la función de esperanza  $\mu(t)$  y la función de autocovarianza  $\gamma(t, s)$  de  $(X_t)_{t \in \mathbb{Z}}$  por

$$\mu(t) := \mathbb{E}[X_t], \quad t \in \mathbb{Z},$$

## 2.1. FUNDAMENTOS DEL ANÁLISIS DE LAS SERIES DE TIEMPO 19

$$\gamma(t, s) := \mathbb{E}[(X_t - \mu(t))(X_s - \mu(s))], \quad t, s \in \mathbb{Z}.$$

Se sigue que la función de autocovarianza satisface que  $\gamma(t, s) = \gamma(s, t)$  y  $\gamma(t, t) = \text{Var}(X_t)$ ,  $\forall t, s \in \mathbb{Z}$ . Generalmente, los procesos que consideraremos serán estacionarios en alguna de las siguientes dos formas.

**Definición 2.1.** *La serie de tiempo  $(X_t)_{t \in \mathbb{Z}}$  es estrictamente estacionaria si*

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{d}{=} (X_{t_1+k}, \dots, X_{t_n+k}),$$

para toda  $t_1, \dots, t_n, k \in \mathbb{Z}$  y para toda  $n \in \mathbb{N}$ .

**Definición 2.2.** *La serie de tiempo  $(X_t)_{t \in \mathbb{Z}}$  es de covarianza estacionaria (o débilmente estacionaria o estacionaria de segundo orden) si los primeros dos momentos existen y satisfacen*

$$\mu(t) = \mu, \quad t \in \mathbb{Z},$$

$$\gamma(t, s) = \gamma(t+k, s+k), \quad t, s, k \in \mathbb{Z}.$$

Ambas definiciones formalizan la noción de que el comportamiento de las series de tiempo es similar en cualquier período que observemos. Una serie de tiempo estrictamente estacionaria con varianza finita es de covarianza estacionaria.

La definición de covarianza estacionaria implica que  $\forall t, s \in \mathbb{Z}$  tenemos que  $\gamma(t-s, 0) = \gamma(t, s) = \gamma(s, t) = \gamma(0, s-t)$ , de tal modo que la covarianza entre  $X_t$  y  $X_s$  sólo depende del valor de  $|s-t|$  que es conocida como el retraso.

**Observación 2.2.** *Para un proceso de covarianza estacionaria, escribimos la función de autocovarianza como una función de una variable*

$$\gamma(h) := \gamma(h, 0), \quad \forall h \in \mathbb{Z}.$$

Nótese que  $\gamma(0) = \text{Var}(X_t)$ ,  $\forall t \in \mathbb{Z}$ .

**Definición 2.3.** *Definimos la función de autocorrelación (ACF)  $\rho(h)$  de un proceso  $(X_t)_{t \in \mathbb{Z}}$  de covarianza estacionaria como*

$$\rho(h) := \rho(X_h, X_0) := \gamma(h)/\gamma(0), \quad \forall h \in \mathbb{Z}.$$

Definiremos ahora otro concepto que nos permitirá construir los modelos de series de tiempo.

**Definición 2.4.** *Decimos que la serie de tiempo  $(X_t)_{t \in \mathbb{Z}}$  es un proceso de ruido blanco si es de covarianza estacionaria con función de autocorrelación*

$$\rho(h) = \begin{cases} 1, & \text{si } h = 0, \\ 0, & \text{si } h \neq 0. \end{cases}$$

Un proceso de ruido blanco centrado tendrá media  $\mu = 0$  y varianza  $\sigma^2 = \text{Var}(X_t)$ , y lo denotaremos como  $\text{WN}(0, \sigma^2)$ .

**Definición 2.5.** *Decimos que la serie de tiempo  $(X_t)_{t \in \mathbb{Z}}$  es un proceso de ruido blanco estricto si es una serie de v.a.i.i.d. con varianza finita.*

Un proceso de ruido blanco estricto centrado tendrá media cero y varianza  $\sigma^2$ , y lo denotaremos como  $\text{SWN}(0, \sigma^2)$ .

Otro concepto que usaremos es el de sucesión de diferencia martingala. Para hablar de este concepto necesitamos suponer que la serie de tiempo  $(X_t)_{t \in \mathbb{Z}}$  es adaptada a alguna filtración  $(\mathcal{F}_t)_{t \in \mathbb{Z}}$  que representa la información acumulada hasta el tiempo  $t$ ; ésta esta definida como  $\mathcal{F}_t = \sigma(\{X_s : s \leq t\})$  y es conocida como la filtración natural.

**Definición 2.6.** *La serie de tiempo  $(X_t)_{t \in \mathbb{Z}}$  es conocida como una sucesión de diferencias martingalas con respecto a la filtración  $(\mathcal{F}_t)_{t \in \mathbb{Z}}$  si  $\mathbb{E}|X_t| < \infty$ ,  $X_t$  es  $\mathcal{F}_t$ -medible (adaptada) y*

$$\mathbb{E}[X_t | \mathcal{F}_{t-1}] = 0, \quad \forall t \in \mathbb{Z}.$$

*La esperanza incondicional del proceso también es cero:*

$$\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_t | \mathcal{F}_{t-1}]] = 0, \quad \forall t \in \mathbb{Z}.$$

*Si  $\mathbb{E}[X_t^2] < \infty \forall t \in \mathbb{Z}$ , entonces la autocovarianza satisface que*

$$\gamma(t, s) = \mathbb{E}[X_t X_s] = \begin{cases} \mathbb{E}[\mathbb{E}[X_t X_s | \mathcal{F}_{s-1}]] = \mathbb{E}[X_t \mathbb{E}[X_s | \mathcal{F}_{s-1}]] = 0, & \text{si } t < s, \\ \mathbb{E}[\mathbb{E}[X_t X_s | \mathcal{F}_{t-1}]] = \mathbb{E}[X_s \mathbb{E}[X_t | \mathcal{F}_{t-1}]] = 0, & \text{si } t > s. \end{cases}$$

*Entonces si una sucesión de diferencia martingala tiene varianza constante  $\forall t \in \mathbb{Z}$  será un proceso de ruido blanco.*

## 2.2. Modelos de series de tiempo ARCH y GARCH

En esta sección daremos las definiciones de los modelos ARCH (*autoregressive conditionally heteroscedastic*) que fueron desarrollados en Engle (1982) y los modelos GARCH (generalización del modelo ARCH) que fueron desarrollados en Bollerslev (1986).

### 2.2.1. Procesos ARCH

**Definición 2.7.** Sea  $(Z_t)_{t \in \mathbb{Z}}$  un  $SWN(0, 1)$ . Diremos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso ARCH( $p$ ) si es estrictamente estacionario y satisface que para toda  $t \in \mathbb{Z}$  y para algún proceso con valores estrictamente positivos  $(\sigma_t)_{t \in \mathbb{Z}}$ , las siguientes ecuaciones

$$X_t = \sigma_t Z_t, \quad (2.1)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2, \quad (2.2)$$

donde  $\alpha_0 > 0$  y  $\alpha_i \geq 0$ ,  $i = 1, \dots, p$ .

Sea  $\mathcal{F}_t = \sigma(\{X_s : s \leq t\})$  la sigma álgebra que representa la historia del proceso hasta el tiempo  $t$ . Es claro que la construcción (2.2) asegura que  $\sigma_t$  es  $\mathcal{F}_{t-1}$  medible. Esto nos permite calcular

$$\mathbb{E}[X_t | \mathcal{F}_{t-1}] = \mathbb{E}[\sigma_t Z_t | \mathcal{F}_{t-1}] = \sigma_t \mathbb{E}[Z_t | \mathcal{F}_{t-1}] = \sigma_t \mathbb{E}[Z_t] = 0. \quad (2.3)$$

Por lo que el proceso ARCH tiene la propiedad de diferencia martingala con respecto a  $(\mathcal{F}_t)_{t \in \mathbb{Z}}$ . Si el proceso es de covarianza estacionaria, es simplemente un ruido blanco.

Si nosotros simplemente asumimos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso de ruido blanco de covarianza estacionaria, entonces  $\mathbb{E}[X_t^2] < \infty$  y

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = \mathbb{E}[\sigma_t^2 Z_t^2 | \mathcal{F}_{t-1}] = \sigma_t^2 \text{Var}(Z_t) = \sigma_t^2.$$

Entonces el modelo tiene una propiedad interesante, que su desviación estándar condicional  $\sigma_t$ , o volatilidad, es una función que cambia continuamente en

los valores cuadrados previos del proceso. Si uno o más de  $|X_{t-1}|, \dots, |X_{t-p}|$  son particularmente grandes, entonces  $X_t$  se obtiene de una distribución con varianza grande, y puede ser ella misma grande; en este sentido el modelo genera la volatilidad de los *clusters*. El nombre ARCH se refiere a esta estructura; el modelo es autoregresivo, pues  $X_t$  es claramente dependiente de  $X_{t-i}$ , y condicionalmente heterocedástico, pues la varianza condicional cambia a cada momento.

La distribución de las innovaciones  $(Z_t)_{t \in \mathbb{Z}}$  puede, en principio, ser cualquier distribución con media cero, y de varianza uno.

### 2.2.2. El modelo ARCH(1)

Usando  $X_t^2 = \sigma_t^2 Z_t^2$  y (2.2) en el caso que  $p = 1$ , deducimos que el proceso cuadrático ARCH(1) satisface

$$X_t^2 = \alpha_0 Z_t^2 + \alpha_1 Z_t^2 X_{t-1}^2. \quad (2.4)$$

Queremos investigar cuando esta ecuación tiene una solución estacionaria expresada en términos de la historia infinita de las innovaciones, es decir, soluciones de la forma  $X_t^2 = f(Z_t, Z_{t-1}, \dots)$ .

El modelo cuadrático ARCH(1) de la ecuación (2.4) es una Ecuación de Recurrencia Estocástica (SRE por sus siglas en inglés) de la forma

$$Y_t = A_t Y_{t-1} + B_t,$$

con  $A_t = \alpha_1 Z_t^2$  y  $B_t = \alpha_0 Z_t^2$ . Se puede inferir que la condición para la solución de la estacionariedad estricta es  $\mathbb{E}[\ln(\alpha_1 Z_t^2)] < 0$ ; y se puede demostrar que es una condición necesaria y suficiente (ver Bougerol y Picard (1992)). Entonces la ecuación (2.4) toma la forma

$$X_t^2 = \alpha_0 \sum_{i=0}^{\infty} \alpha_1^i \prod_{j=0}^i Z_{t-j}^2. \quad (2.5)$$

Si las innovaciones  $(Z_t)$  son normal estándar, entonces la condición para una solución estrictamente estacionaria es aproximadamente  $\alpha_1 < 3.562$ ; y si las  $(Z_t)$  son innovaciones  $t$  con cuatro grados de libertad y varianza 1, la condición es  $\alpha_1 < 5.437$ .

La estacionariedad estricta depende en la distribución de las innovaciones, aunque la covarianza estacionaria no; la condición suficiente y necesaria para la covarianza estacionaria siempre es  $\alpha_1 < 1$ , que ahora probaremos.

**Proposición 2.1.** *El proceso ARCH(1) es un proceso de ruido blanco de covarianza estacionaria si y sólo si  $\alpha_1 < 1$ . La varianza del proceso de covarianza estacionaria está dado por  $\alpha_0/(1 - \alpha_1)$ .*

*Demostración.* Supongamos que el proceso es de covarianza estacionaria y que  $\mathbb{E}[Z_t^2] = 1$  tal que

$$\sigma^2 = \mathbb{E}[X_t^2] = \alpha_0 + \alpha_1 \mathbb{E}[X_{t-1}^2] = \alpha_0 + \alpha_1 \sigma^2.$$

Claramente,  $\sigma^2 = \alpha_0/(1 - \alpha_1)$  y tenemos que  $\alpha_1 < 1$ .

Inversamente, si  $\alpha_1 < 1$ , entonces por la desigualdad de Jensen,

$$\mathbb{E}[\ln(\alpha_1 Z_t^2)] \leq \ln(\mathbb{E}[\alpha_1 Z_t^2]) = \ln(\alpha_1) < 0,$$

usando (2.5)

$$\mathbb{E}[X_t^2] = \alpha_0 \sum_{i=0}^{\infty} \alpha_1^i = \frac{\alpha_0}{1 - \alpha_1}.$$

□

### 2.2.3. Procesos GARCH

**Definición 2.8.** *Sea  $(Z_t)_{t \in \mathbb{Z}}$  un SWN(0, 1). Diremos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso GARCH(p, q) si es estrictamente estacionario y satisface que para toda  $t \in \mathbb{Z}$  y para algún proceso con valores estrictamente positivos  $(\sigma_t)_{t \in \mathbb{Z}}$ , las siguientes ecuaciones*

$$X_t = \sigma_t Z_t, \tag{2.6}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (2.7)$$

donde  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, \dots, p$  y  $\beta_j \geq 0$ ,  $j = 1, \dots, q$ .

Los procesos GARCH son generalizaciones de procesos ARCH, en el sentido que la volatilidad al cuadrado,  $\sigma_t^2$ , depende de volatilidades cuadradas previas, así como valores previos al cuadrado del proceso  $(X_t)_{t \in \mathbb{Z}}$ .

En la práctica son más usados los modelos GARCH de un orden bajo dado que son fáciles de manejar y al parecer se ajustan bien a los modelos de series financieras. En estos modelos los períodos de alta volatilidad tienden a ser persistentes; se pueden obtener los mismos resultados con modelos ARCH de un orden bastante elevado, pero los modelos GARCH de orden bajo logran este efecto y de una manera parsimoniosa.

Ahora si en (2.7) nos tomamos  $p = 1 = q$ , entonces tenemos un modelo GARCH(1,1)

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

y si elevamos al cuadrado (2.6) y sustituimos

$$\sigma_t^2 = \alpha_0 + \alpha_1 (\sigma_{t-1}^2 Z_{t-1}^2) + \beta_1 \sigma_{t-1}^2 = \alpha_0 + \sigma_{t-1}^2 (\alpha_1 Z_{t-1}^2 + \beta_1).$$

Nuevamente tenemos una ecuación de recurrencia estocástica, pero ahora la condición de estacionariedad se transforma en

$$\mathbb{E}[\ln(\alpha_0 Z_t^2 + \beta_1)] < 0,$$

y la solución general se vuelve

$$\sigma_t^2 = \alpha_0 + \alpha_0 \sum_{i=1}^{\infty} \prod_{j=1}^i (\alpha_1 Z_{t-j}^2 + \beta_1). \quad (2.8)$$

Si  $(\sigma_t^2)_{t \in \mathbb{Z}}$  es un proceso estrictamente estacionario, entonces también  $(X_t)_{t \in \mathbb{Z}}$  lo es, porque  $X_t = \sigma_t Z_t$ , dado que  $(Z_t)_{t \in \mathbb{Z}}$  es simplemente un ruido blanco estricto. La solución del proceso GARCH(1,1) es entonces

$$X_t = Z_t \sqrt{\alpha_0 \left( 1 + \sum_{i=1}^{\infty} \prod_{j=1}^i (\alpha_1 Z_{t-j}^2 + \beta_1) \right)}. \quad (2.9)$$

**Proposición 2.2.** *El proceso GARCH(1,1) es un proceso de ruido blanco con covarianza estacionaria si y sólo si  $\alpha_1 + \beta_1 < 1$ . La varianza del proceso de covarianza estacionaria está dada por  $\alpha_0 / (1 - \alpha_1 - \beta_1)$ .*

*Demostración.* Al igual que en la Proposición 2.1, si suponemos que  $\mathbb{E}[Z_t^2] = 1$ , entonces

$$\sigma^2 = \mathbb{E}[X_t^2] = \alpha_0 + \alpha_1 \mathbb{E}[X_{t-1}^2] + \beta_1 \mathbb{E}[X_{t-1}^2] = \alpha_0 + \alpha_1 \sigma^2 + \beta_1 \sigma^2.$$

Claramente,  $\sigma^2 = \alpha_0 / (1 - \alpha_1 - \beta_1)$  y tenemos que  $\alpha_1 + \beta_1 < 1$ .

Inversamente, si  $\alpha_1 + \beta_1 < 1$ , entonces por la desigualdad de Jensen,

$$\mathbb{E}[\ln(\alpha_1 Z_t^2 + \beta_1)] \leq \ln(\mathbb{E}[\alpha_1 Z_t^2 + \beta_1]) = \ln(\alpha_1 + \beta_1) < 0,$$

usando (2.9)

$$\mathbb{E}[X_t^2] = \alpha_0 \left( 1 + \sum_{i=1}^{\infty} (\alpha_1 + \beta_1)^i \right) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

□

#### 2.2.4. Ajuste de modelos GARCH

Para el ajuste de los modelos GARCH usualmente se emplea la función de verosimilitud, por simplicidad lo haremos para los modelos de primer orden, es decir, para el ARCH(1) y para el GARCH(1, 1).

Suponiendo que tenemos un total de  $n + 1$  datos correspondientes a  $X_0, \dots, X_n$ . Podemos escribir la densidad conjunto de las variables como

$$f_{X_0, \dots, X_n}(x_0, \dots, x_n) = f_{X_0}(x_0) \prod_{t=1}^n f_{X_t|X_{t-1}, \dots, X_0}(x_t|x_{t-1}, \dots, x_0). \quad (2.10)$$

Para un proceso ARCH(1), las densidades condicionales dependen sólo del valor pasado de  $\sigma_t$  o equivalentemente de  $X_{t-1}$ . Así, la densidad puede ser calculada como

$$f_{X_t|X_{t-1}, \dots, X_0}(x_t|x_{t-1}, \dots, x_0) = f_{X_t|X_{t-1}}(x_t|x_{t-1}) = \frac{1}{\sigma_t} g\left(\frac{x_t}{\sigma_t}\right),$$

donde  $\sigma_t = (\alpha_0 + \alpha_1 x_{t-1}^2)^{1/2}$  y  $g(z)$  denota la densidad marginal de las innovaciones  $(Z_t)_{t \in \mathbb{Z}}$ , las cuales deberán tener media cero y varianza uno. Se suelen elegir distribuciones normal estándar o distribuciones  $t$  escaladas para que tengan varianza uno.

Sin embargo, no conocemos la densidad de  $f_{X_0}$ . Una posible solución es construir la verosimilitud condicional dado  $X_0$ , la cual es calculada como

$$f_{X_1, \dots, X_n|X_0}(x_1, \dots, x_n|x_0) = \prod_{t=1}^n f_{X_t|X_{t-1}, \dots, X_0}(x_t|x_{t-1}, \dots, x_0).$$

Para el modelo ARCH(1) se sigue que

$$L(\alpha_0, \alpha_1; \mathbf{X}) = f_{X_1, \dots, X_n|X_0}(X_1, \dots, X_n|X_0) = \prod_{t=1}^n \frac{1}{\sigma_t} g\left(\frac{X_t}{\sigma_t}\right),$$

donde  $\sigma_t = (\alpha_0 + \alpha_1 X_{t-1}^2)^{1/2}$ .

Para el GARCH(1, 1), construimos la densidad conjunta de  $X_1, \dots, X_n$  condicionada a los valores de  $X_0$  y de  $\sigma_0$ , la cual es

$$f_{X_1, \dots, X_n|X_0, \sigma_0}(x_1, \dots, x_n|x_0, \sigma_0) = \prod_{t=1}^n f_{X_t|X_{t-1}, \dots, X_0, \sigma_0}(x_t|x_{t-1}, \dots, x_0, \sigma_0).$$

Las densidades condicionales  $f_{X_t|X_{t-1}, \dots, X_0, \sigma_0}$  dependen solamente del valor pasado de  $\sigma_t$ , el cual está dado de manera recursiva por  $\sigma_0, X_0, \dots, X_{t-1}$  usando  $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2$ . Esto nos da la verosimilitud condicional

$$L(\alpha_0, \alpha_1, \beta; \mathbf{X}) = \prod_{t=1}^n \frac{1}{\sigma_t} g\left(\frac{X_t}{\sigma_t}\right), \quad \sigma_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2}.$$

De nuevo, no tenemos un valor para  $\sigma_0^2$ , así que podemos tomar un valor inicial como la varianza muestral de  $X_1, \dots, X_n$  o tomándola igual a cero.

Consideremos ahora a una log-verosimilitud de la forma

$$\ln L(\theta; \mathbf{X}) = \sum_{t=1}^n l_t(\theta),$$

donde  $l_t$  denota la contribución a log-verosimilitud de la  $t$ -ésima observación. El estimador máximo verosímil  $\hat{\theta}$  maximiza la log-verosimilitud (condicional) anterior y siendo un máximo local, resuelve la ecuación de verosimilitud

$$\frac{\partial}{\partial \theta} \ln L(\theta; \mathbf{X}) = \sum_{t=1}^n \frac{\partial l_t(\theta)}{\partial \theta} = \mathbf{0},$$

donde la parte izquierda es conocida como el *vector de scores* de la verosimilitud condicional. Este tipo de ecuaciones se resuelven con métodos numéricos.

Para saber si el modelo que hemos ajustado es correcto, se revisan los residuales. Los residuales estandarizados deberán comportarse como una realización de un proceso SWN. Suponiendo que esta hipótesis no es rechazada, por lo que la dinámica ha sido capturada de forma satisfactoria, la validez de la distribución puede ser investigada a través del uso del Gráfico *QQ*.

## 2.3. Media Móvil con Ponderación Exponencial

Si nuestros datos de rendimientos  $X_{t-n+1}, \dots, X_t$  siguen algún tipo de modelo de serie de tiempo en el que se define una volatilidad, y quisiéramos

el pronóstico de nuestras observaciones al cuadrado, tendríamos

$$X_{t+1}^2 = \sum_{i=0}^{n-1} \lambda(1-\lambda)^i X_{t-i}^2, \quad 0 < \lambda < 1.$$

De esta manera ponderamos los datos del más reciente al más antiguo con una sucesión de pesos exponenciales decrecientes que suman 1. Entonces, podemos calcular

$$\begin{aligned} X_{t+1}^2 &= \sum_{i=0}^{n-1} \lambda(1-\lambda)^i X_{t-i}^2 = \lambda X_t^2 + (1-\lambda) \sum_{j=0}^{n-2} \lambda(1-\lambda)^j X_{t-1-j}^2 \\ &= \lambda X_t^2 + (1-\lambda) P_{t-1} X_t^2, \end{aligned}$$

de manera que la predicción al tiempo  $t + 1$  se obtiene recursivamente de la predicción al tiempo  $t$ . La elección de  $\lambda$  es subjetiva; entre mayor sea el valor, mayor será el peso de la observación más reciente.

Como la esperanza de  $X_{t+1}^2$  y  $\sigma_{t+1}^2$  coinciden, podemos definir la ecuación anterior para la volatilidad, es decir

$$\hat{\sigma}_{t+1}^2 = \lambda X_t^2 + (1-\lambda) \hat{\sigma}_t^2. \quad (2.11)$$

Esta es la idea esencial del enfoque de las Medias Móviles con Ponderaciones Exponenciales para la predicción de la volatilidad. Este modelo es mejor conocido por EWMA debido a su nombre en inglés (*Exponentially Weighted Moving Average*), y se usa principalmente cuando queremos predecir la volatilidad y nuestro modelo GARCH no cumple que  $\sum \alpha_i + \sum \beta_j < 1$ .

## 2.4. Fundamentos de las series de tiempo multivariadas

Abordaremos ahora las propiedades y resultados esenciales de las series de tiempo multivariadas, y a continuación la estructura general de los mode-

los Multivariados GARCH.

Una serie de tiempo multivariada para múltiples factores de riesgo es un proceso estocástico  $(X_t)_{t \in \mathbb{Z}}$ , es decir, una familia de vectores aleatorios, indexados en los enteros y definidos en un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ .

**Definición 2.9.** *Asumiendo que los primeros dos momentos de una serie de tiempo multivariada existen, definimos la función de medias  $\mu(t)$  y la función de matriz de covarianzas  $\Gamma(t, s)$  de  $(X_t)_{t \in \mathbb{Z}}$  como*

$$\mu(t) := \mathbb{E}[X_t], \quad t \in \mathbb{Z},$$

$$\Gamma(t, s) := \mathbb{E}[(X_t - \mu(t))(X_s - \mu(s))'], \quad t, s \in \mathbb{Z},$$

donde  $\mathbb{E}[X_t]$  es el vector de las esperanzas de  $X_t$  y  $\Gamma(t, s)$  es la matriz de covarianzas entre  $X_t$  y  $X_s$ .

De manera análoga, en el caso univariado tenemos que  $\Gamma(t, t) = \text{Var}(X_t)$ . Observemos que en este caso los elementos  $\gamma_{ij}(t, s)$  de  $\Gamma(t, s)$  satisfacen que

$$\gamma_{ij}(t, s) = \text{Cov}(X_{t,i}, X_{s,j}) = \text{Cov}(X_{s,j}, X_{t,i}) = \gamma_{ji}(s, t),$$

es claro que  $\Gamma(t, s) = \Gamma(s, t)' \forall t, s \in \mathbb{Z}$ . Sin embargo, la matriz  $\Gamma$  no necesariamente es simétrica, así que en general  $\Gamma(t, s) \neq \Gamma(s, t)$ .

Los modelos multivariados que consideraremos serán estacionarios en al menos una de las siguientes formas.

**Definición 2.10.** *La serie de tiempo multivariada  $(X_t)_{t \in \mathbb{Z}}$  es estrictamente estacionaria si*

$$(X'_{t_1}, \dots, X'_{t_n}) \stackrel{d}{=} (X'_{t_1+k}, \dots, X'_{t_n+k}),$$

para toda  $t_1, \dots, t_n, k \in \mathbb{Z}$  y para toda  $n \in \mathbb{N}$ .

**Definición 2.11.** *La serie de tiempo multivariada  $(X_t)_{t \in \mathbb{Z}}$  es de covarianza estacionaria (débilmente estacionaria o estacionaria de segundo orden) si los primeros dos momentos existen y satisfacen que*

$$\mu(t) = \mu, \quad t \in \mathbb{Z},$$

$$\Gamma(t, s) = \Gamma(t+k, s+k), \quad t, s, k \in \mathbb{Z}.$$

**Observación 2.3.** *Una serie de tiempo estrictamente estacionaria con matriz de covarianzas finita, es de covarianza estacionaria.*

La definición de covarianza estacionaria implica que para toda  $s, t$  tenemos que  $\Gamma(t - s, 0) = \Gamma(t, s)$ , así que la covarianza entre  $X_t$  y  $X_s$  depende únicamente de la separación temporal  $t - s$ , que es conocido como el retraso. A diferencia del caso univariado, el signo del retraso sí importa.

Para un proceso multivariado de covarianza estacionaria, escribimos la función de la matriz de covarianza como una función de una única variable  $\Gamma(h) := \Gamma(h, 0)$ ,  $\forall h \in \mathbb{Z}$ , notemos que  $\Gamma(0) = \text{Var}(X_t)$ ,  $\forall t \in \mathbb{Z}$ . Con estos elementos podemos definir la función de la matriz de correlación de un proceso de covarianza estacionaria.

**Definición 2.12.** *Definiremos el operador  $\Delta := \Delta(\Gamma(0))$ , donde el operador  $\Delta(\Sigma) := \text{diag}(\sqrt{\sigma_{11}}, \dots, \sqrt{\sigma_{dd}})$  para alguna matriz  $\Sigma$ .*

**Definición 2.13.** *Definimos la función de la matriz de correlación  $P(h)$  de una serie de tiempo multivariada de covarianza estacionaria  $(X_t)_{t \in \mathbb{Z}}$  como*

$$P(h) := \Delta^{-1} \Gamma(h) \Delta^{-1}, \quad \forall h \in \mathbb{Z}.$$

Como en el caso univariado, los procesos multivariados de ruido blanco son procesos que nos ayudan a crear modelos de series de tiempo.

**Definición 2.14.** *El proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso multivariado de ruido blanco si es de covarianza estacionaria con función de matriz de correlación dada por*

$$P(h) = \begin{cases} P, & h = 0, \\ 0, & h \neq 0, \end{cases}$$

*para alguna matriz de correlación positiva definida  $P$ .*

Un proceso multivariado de ruido blanco con media cero y matriz de covarianzas  $\Sigma$ , será denotado por  $\text{WN}(0, \Sigma)$ .

**Definición 2.15.** *El proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso multivariado de ruido blanco estricto si es una serie de vectores aleatorios i.i.d. con matriz de covarianzas finita.*

Un proceso multivariado de ruido blanco estricto con media cero y matriz de covarianzas  $\Sigma$ , será denotado por SWN(0,  $\Sigma$ ).

**Definición 2.16.** *El proceso multivariado  $(X_t)_{t \in \mathbb{Z}}$  tiene la propiedad de diferencia martingala multivariada con respecto a la filtración  $(\mathcal{F}_t)$ , si cumple que  $\mathbb{E}[|X_t|] < \infty$  y*

$$\mathbb{E}[X_t | \mathcal{F}_{t-1}] = 0, \quad \forall t \in \mathbb{Z}.$$

**Observación 2.4.** *En la práctica, la función de la matriz de covarianza  $\Gamma(h)$  y la función de la matriz de correlación  $P(h)$ , son reemplazadas por sus correspondientes estimadores, es decir, por la función de la matriz de covarianza muestral*

$$\hat{\Gamma}(h) := \frac{1}{n} \sum_{t=1}^{n-h} (X_{t-h} - \bar{X})(X_t - \bar{X})', \quad 0 \leq h < n,$$

y por la función de la matriz de correlación muestral

$$\hat{P}(h) := \hat{\Delta}^{-1} \hat{\Gamma}(h) \hat{\Delta}^{-1}, \quad 0 \leq h < n,$$

donde  $\bar{X} := \sum_{t=1}^n X_t/n$  es la media muestral que estima a  $\mu$ , y  $\hat{\Delta} := \Delta(\hat{\Gamma}(0))$ .

## 2.5. Procesos GARCH Multivariados

Es bien sabido que la volatilidad de los datos financieros a menudo varía con el tiempo y tiende a agruparse en períodos, es decir, alta volatilidad es seguida generalmente por alta volatilidad y baja volatilidad es seguida por baja volatilidad. Los modelos GARCH y sus extensiones han demostrado ser capaces de capturar la agrupación de la volatilidad y predecirla, además de poder analizar el movimiento de los rendimientos financieros.

La idea básica de extender los modelos GARCH univariados a modelos GARCH Multivariados, es que estos últimos son importantes para predecir la dependencia de los movimientos conjuntos de los rendimientos de los activos de una cartera, y al realizar un modelo multivariado tendríamos un modelo mucho más confiable que usar varios modelos univariados por separado.

Al modelar con un MGARCH, se parametriza la matriz de covarianzas, ya que es factible en términos de estimación y se tiene una mínima pérdida de generalidad. La mayoría de los modelos MGARCH existentes, parametrizan la matriz de covarianzas con un gran número de parámetros que son difíciles de estimar y usualmente nos llevan a problemas de convergencia en la estimación de los algoritmos.

La elección del modelo MGARCH se determina a menudo por consideraciones prácticas, ya que por un lado, el modelo debe ser lo suficientemente flexible para indicar la dinámica de las varianzas y covarianzas condicionales; y por otro lado, el número de parámetros en un modelo MGARCH aumentará rápidamente conforme la dimensión del modelo aumente; por lo que se usa el criterio de parsimonia para simplificar la estimación del mismo y facilitar la interpretación de sus parámetros. Sin embargo, parsimonia a menudo significa simplificación, y los modelos con muy pocos parámetros no son capaces de capturar las dinámicas relevantes en la estructura de la matriz de covarianzas.

### 2.5.1. Estructura general de los modelos

**Definición 2.17.** Sea  $(Z_t)_{t \in \mathbb{Z}}$  un  $SWN(0, I_d)$ . Se dice que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso GARCH Multivariado si es estrictamente estacionario y satisface la ecuación de la forma

$$X_t = \Sigma_t^{1/2} Z_t, \quad t \in \mathbb{Z}, \quad (2.12)$$

donde  $\Sigma_t^{1/2} \in \mathbb{R}^{d \times d}$  es el factor de Cholesky de una matriz positiva definida  $\Sigma_t$ , que es medible con respecto a  $\mathcal{F}_{t-1} = \sigma(\{X_s : s \leq t-1\})$  la historia del proceso hasta  $t-1$ .

Se puede observar que un proceso de covarianza estacionaria de este tipo tiene la propiedad de diferencia martingala multivariada

$$\mathbb{E}[X_t | \mathcal{F}_{t-1}] = \mathbb{E}[\Sigma_t^{1/2} Z_t | \mathcal{F}_{t-1}] = \Sigma_t^{1/2} \mathbb{E}[Z_t] = 0,$$

y por lo tanto debe ser un proceso de ruido blanco. Más aún,  $\Sigma_t$  será la matriz de covarianzas condicional ya que

$$\text{Cov}(X_t | \mathcal{F}_{t-1}) = \mathbb{E}[X_t X_t' | \mathcal{F}_{t-1}] = \Sigma_t^{1/2} \mathbb{E}[Z_t Z_t'] \left(\Sigma_t^{1/2}\right)' = \Sigma_t^{1/2} \left(\Sigma_t^{1/2}\right)' = \Sigma_t.$$

La matriz de covarianzas condicional  $\Sigma_t$  en un modelo GARCH Multivariado, corresponde a la volatilidad al cuadrado  $\sigma_t^2$  en un modelo GARCH univariado.

Denotaremos a los elementos de  $\Sigma_t$  por  $\sigma_{t,ij}$ , y haremos uso de la notación  $\sigma_{t,i} = \sqrt{\sigma_{t,ii}}$  para referirnos a la desviación estándar condicional (o volatilidad) de la  $i$ -ésima componente de la serie  $(X_{t,i})_{t \in \mathbb{Z}}$ .

Recordemos que podemos escribir a  $\Sigma_t = \Delta_t P_t \Delta_t$ , donde

$$\Delta_t = \Delta(\Sigma_t) = \text{diag}(\sigma_{t,1}, \dots, \sigma_{t,d}), \quad P_t = \wp(\Sigma_t),$$

donde el operador  $\wp(\Sigma_t)$ , extrae la matriz de correlación de la matriz  $\Sigma$ .

La matriz diagonal  $\Delta_t$  será conocida como la matriz de volatilidad y  $P_t$  como la matriz de correlación condicional. Para construir modelos GARCH Multivariados necesitamos especificar la dependencia de  $\Sigma_t$  (o de  $\Delta_t$  y  $P_t$ ), de tal manera que  $\Sigma_t$  siempre sea simétrica y positiva semidefinida

**Definición 2.18.** *La matriz de covarianzas incondicional  $\Sigma$ , está dada por*

$$\Sigma = \text{Cov}(X_t) = \mathbb{E}[\text{Cov}(X_t | \mathcal{F}_{t-1})] + \text{Cov}(\mathbb{E}[X_t | \mathcal{F}_{t-1}]) = \mathbb{E}[\Sigma_t],$$

a partir de la cual se puede calcular la matriz de correlación incondicional  $P$ , cuyos elementos son

$$\rho_{ij} = \frac{\mathbb{E}[\sigma_{t,ij}]}{\sqrt{\mathbb{E}[\sigma_{t,ii}]\mathbb{E}[\sigma_{t,jj}]}} = \frac{\mathbb{E}[\rho_{t,ij}\sigma_{t,i}\sigma_{t,j}]}{\sqrt{\mathbb{E}[\sigma_{t,i}^2]\mathbb{E}[\sigma_{t,j}^2]}}$$

los cuales son difíciles de evaluar, y no es simplemente la esperanza de la matriz de correlación condicional.

**Observación 2.5.** *En la práctica, las innovaciones son generalmente una distribución Gaussiana multivariada ( $Z_t \sim N_d(0, I_d)$ ), o cuando se trabaja con rendimientos diarios, es mejor utilizar una distribución esférica multivariada  $t$  ( $Z_t \sim t_d(\nu, 0, (\nu - 2)I_d/\nu)$ ).*

A continuación, veremos algunos de los modelos GARCH Multivariados más importantes, las condiciones necesarias para asegurar que la matriz de covarianzas  $\Sigma_t$  sea positiva definida y analizaremos los problemas que surgen cuando aumenta la dimensionalidad.

### 2.5.2. Modelos de correlación condicional

Estos modelos se centran en especificar la matriz de correlación condicional  $P_t$  y permiten que las volatilidades sean descritas por modelos GARCH univariados.

#### CCC-GARCH (GARCH de Correlación Condicional Constante)

**Definición 2.19.** Diremos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso CCC-GARCH si es un proceso con la estructura general de la Definición 2.17 y su matriz de covarianzas condicional es de la forma

$$\Sigma_t = \Delta_t P_c \Delta_t,$$

donde  $P_c$  es una matriz de correlación constante positiva definida y  $\Delta_t$  es la matriz diagonal de la volatilidad cuyos elementos  $\sigma_{t,k}$  satisfacen

$$\sigma_{t,k}^2 = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} X_{t-i,k}^2 + \sum_{j=1}^{q_k} \beta_{kj} \sigma_{t-j,k}^2, \quad k = 1, \dots, d,$$

donde  $\alpha_{k0} > 0$ ;  $\alpha_{ki} \geq 0$ ,  $i = 1, \dots, p_k$ ;  $\beta_{kj} \geq 0$ ,  $j = 1, \dots, q_k$ .

**Proposición 2.3.** El modelo CCC-GARCH está bien definido, en el sentido que  $\Sigma_t$  es positiva definida c.s. para toda  $t$ . Además, es de covarianza estacionaria si y solo si  $\sum_{i=1}^{p_k} \alpha_{ki} + \sum_{j=1}^{q_k} \beta_{kj} < 1$  para  $k = 1, \dots, d$ .

*Demostración.* Para algún vector  $v \neq 0$  en  $\mathbb{R}^d$  tenemos que

$$v' \Sigma_t v = (\Delta_t v)' P_c (\Delta_t v) > 0,$$

dado que  $P_c$  es positiva definida y la positividad estricta de los procesos de volatilidad individuales nos asegura que  $\Delta_t v \neq 0 \forall t \in \mathbb{Z}$ . Si el proceso  $(X_t)_{t \in \mathbb{Z}}$  es de covarianza estacionaria, entonces cada componente de la serie  $(X_{t,k})_{t \in \mathbb{Z}}$

es un proceso GARCH de covarianza estacionaria para el cual  $\sum_{i=1}^{p_k} \alpha_{ki} + \sum_{j=1}^{q_k} \beta_{kj} < 1$  es una condición necesaria y suficiente por la Proposición 2.2. Inversamente, si las componentes de la serie son de covarianza estacionaria, entonces para toda  $i$  y  $j$ , la desigualdad de Cauchy-Schwarz implica que

$$\sigma_{ij} = \mathbb{E}[\sigma_{t,ij}] = \rho_{ij} \mathbb{E}[\sigma_{t,i}\sigma_{t,j}] \leq \rho_{ij} \sqrt{\mathbb{E}[\sigma_{t,i}^2]} \sqrt{\mathbb{E}[\sigma_{t,j}^2]} < \infty.$$

Como  $(X_t)_{t \in \mathbb{Z}}$  es una diferencia martingala multivariada con segundos momentos finitos y no dependientes del tiempo  $\sigma_{ij}$ , entonces es un ruido blanco de covarianza estacionaria.  $\square$

El modelo CCC-GARCH se usa generalmente como un modelo inicial para proceder luego con modelos más complejos. Una crítica a este modelo es el hecho de suponer la constancia de la matriz de correlación, porque es una característica poco realista y el impacto de las noticias sobre los mercados financieros requiere modelos que permitan una evolución dinámica de la matriz de correlación y de las volatilidades.

**Definición 2.20.** *Para cualquier proceso multivariado  $(X_t)_{t \in \mathbb{Z}}$ , definimos el proceso de desvolatilización como*

$$Y_t := \Delta_t^{-1} X_t,$$

donde  $\Delta_t$  es la matriz diagonal de las volatilidades.

Al ajustar un modelo CCC-GARCH, es necesario estimar primero las volatilidades individuales de las componentes de la serie con procesos GARCH univariados. Luego estimamos el proceso de desvolatilización tomando  $\hat{Y}_t = \hat{\Delta}_t^{-1} X_t$ , donde  $\hat{\Delta}_t^{-1}$  es el estimador de  $\Delta_t$ . Si el modelo CCC-GARCH es adecuado, entonces  $\hat{Y}_t$  debe comportarse como una realización de un proceso SWN(0,  $P_c$ ), y entonces la matriz de correlación condicional  $P_c$  puede ser estimada de los residuales estandarizados.

### DCC-GARCH (GARCH de Correlación Condicional Dinámica)

Este modelo generaliza el modelo CCC-GARCH al permitir que las correlaciones condicionales evolucionen dinámicamente de acuerdo a un esquema relativamente parsimonioso, pero es construido de tal forma que se permita

estimar con modelos GARCH univariados. Su análisis se vuelve complejo por el uso del operador de extracción de la matriz de correlación  $\wp$  en su definición.

**Definición 2.21.** *Diremos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso DCC-GARCH si es un proceso con la estructura general de la Definición 2.17; las volatilidades que comprenden la matriz  $\Delta_t$  siguen especificaciones GARCH univariadas, es decir*

$$\sigma_{t,k}^2 = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} X_{t-i,k}^2 + \sum_{j=1}^{q_k} \beta_{kj} \sigma_{t-j,k}^2, \quad k = 1, \dots, d,$$

donde  $\alpha_{k0} > 0$ ;  $\alpha_{ki} \geq 0$ ,  $i = 1, \dots, p_k$ ;  $\beta_{kj} \geq 0$ ,  $j = 1, \dots, q_k$ ; y las matrices de correlación condicional satisfacen para toda  $t \in \mathbb{Z}$ , las siguientes ecuaciones

$$P_t = \wp \left( \left( 1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j \right) P_c + \sum_{i=1}^p \alpha_i Y_{t-i} Y'_{t-i} + \sum_{j=1}^q \beta_j P_{t-j} \right), \quad (2.13)$$

donde  $P_c$  es una matriz de correlación positiva definida;  $\wp$  es el operador que extrae la matriz de correlación;  $Y_t = \Delta_t^{-1} X_t$  denota el proceso de desvolatilización; y los coeficientes satisfacen que  $\alpha_i \geq 0$ ,  $\forall i = 1, \dots, p$ ;  $\beta_j \geq 0$ ,  $\forall j = 1, \dots, q$ ; y  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ .

Observemos que si todos los coeficientes  $\alpha_i$  y  $\beta_j$  son cero, entonces el modelo se reduce al CCC-GARCH, porque para toda  $t \in \mathbb{Z}$   $P_t = P_c$ . Notemos que la ecuación (2.13) preserva la propiedad de ser positiva definida. Si definimos

$$Q_t := \left( 1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j \right) P_c + \sum_{i=1}^p \alpha_i Y_{t-i} Y'_{t-i} + \sum_{j=1}^q \beta_j P_{t-j},$$

y suponemos que  $P_{t-q}, \dots, P_{t-1}$  son positivas definidas, entonces se sigue que para algún vector  $v \neq 0$  en  $\mathbb{R}^d$ , tenemos

$$v' Q_t v = \left( 1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j \right) v' P_c v + \sum_{i=1}^p \alpha_i v' Y_{t-i} Y'_{t-i} v + \sum_{j=1}^q \beta_j v' P_{t-j} v > 0,$$

pues el primer término es estrictamente positivo y el segundo y tercero son no negativos. Por lo que si  $Q_t$  es positiva definida, entonces  $P_t$  también lo es.

Al ajustar un modelo DCC-GARCH, es necesario ajustar primero modelos GARCH univariados a las componentes de la serie para estimar la matriz de volatilidad  $\Delta_t$ . Enseguida estimamos el proceso de desvolatilización tomando  $\hat{Y}_t = \hat{\Delta}_t X_t$ , y, posteriormente estimamos  $P_c$  tomando la matriz de correlación muestral de los datos desvolatilizados. Para finalizar, estimamos los parámetros  $\alpha_i$  y  $\beta_j$  mediante el ajuste de un modelo con la estructura  $Y_t = P_t^{1/2} Z_t$  a los datos desvolatilizados

### 2.5.3. Modelos de covarianza condicional

Estos modelos especifican explícitamente una estructura dinámica para la matriz de covarianzas  $\Sigma_t$ , y no están diseñados para las estimaciones de modelos GARCH univariados.

#### El modelo VEC y DVEC

**Definición 2.22.** Diremos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso VEC si es un proceso con la estructura general de la Definición 2.17, y las dinámicas de la matriz de covarianzas condicional  $\Sigma_t$  están dadas por las ecuaciones

$$\text{vech}(\Sigma_t) = a_0 + \sum_{i=1}^p \bar{A}_i \text{vech}(X_{t-i} X'_{t-i}) + \sum_{j=1}^q \bar{B}_j \text{vech}(\Sigma_{t-j}), \quad (2.14)$$

para algún vector  $a_0 \in \mathbb{R}^{d(d+1)/2}$  y matrices  $\bar{A}_i$  y  $\bar{B}_j$  en  $\mathbb{R}^{(d(d+1)/2) \times (d(d+1)/2)}$ .

En esta definición “vech” denota el operador que apila las columnas del triángulo inferior de una matriz simétrica en un único vector columna de longitud  $d(d+1)/2$ .

En esta forma general, el modelo tiene  $(1 + (p+q)d(d+1)/2)d(d+1)/2$  parámetros, y estos crecen rápidamente conforme la dimensión aumente, por ejemplo un modelo de tres variables tiene 78 parámetros. Para simplificar este problema, se suele restringir que  $\bar{A}_i$  para  $i = 1, \dots, p$  y  $\bar{B}_j$  para  $j = 1, \dots, q$  sean matrices diagonales, esto nos da el modelo VEC diagonal o el modelo DVEC. Este caso puede ser escrito en términos del *producto de Hadamard*, que se denota con el símbolo “ $\circ$ ”, y es una forma de multiplicar matrices del

mismo tamaño que se realiza elemento a elemento.

**Definición 2.23.** Diremos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso DVEC si es un proceso con la estructura general de la Definición 2.17, y las dinámicas de la matriz de covarianzas condicional  $\Sigma_t$  están dadas por las ecuaciones

$$\Sigma_t = A_0 + \sum_{i=1}^p A_i \circ (X_{t-i} X_{t-i}') + \sum_{j=1}^q B_j \circ \Sigma_{t-j}, \quad (2.15)$$

donde  $A_0$ ,  $A_i$  para  $i = 1, \dots, p$  y  $B_j$  para  $j = 1, \dots, q$  deben ser todas matrices simétricas pertenecientes a  $\mathbb{R}^{d \times d}$ , tal que  $A_0$  tenga elementos diagonales positivos y todas las otras matrices tengan elementos diagonales no negativos.

Una condición suficiente para que  $\Sigma_t$  sea positiva definida c.s. es que  $A_0$  sea positiva definida y las matrices  $A_1, \dots, A_p, B_1, \dots, B_q$  sean todas positivas semidefinidas. Para lograr esta condición, podemos restringir que todas nuestras matrices tengan una descomposición de Cholesky, es decir, podemos parametrizar el modelo en términos de las matrices triangulares de la descomposición de Cholesky  $A_0^{1/2}$ ,  $A_i^{1/2}$  y  $B_j^{1/2}$  que satisfacen

$$A_0 = A_0^{1/2} \left( A_0^{1/2} \right)', \quad A_i = A_i^{1/2} \left( A_i^{1/2} \right)', \quad B_j = B_j^{1/2} \left( B_j^{1/2} \right)'. \quad (2.16)$$

Como la condición sólo pide que  $A_1, \dots, A_p, B_1, \dots, B_q$  sean positivas semidefinidas, podemos considerar parametrizaciones aún más simples, tales como

$$A_0 = A_0^{1/2} \left( A_0^{1/2} \right)', \quad A_i = a_i a_i', \quad B_j = b_j b_j', \quad (2.17)$$

donde  $a_i$  y  $b_j$  son vectores en  $\mathbb{R}^d$ . Inclusive, podemos pedir algo más simple y que todavía satisfaga la condición, ésta es

$$A_0 = A_0^{1/2} \left( A_0^{1/2} \right)', \quad A_i = a_i I_d, \quad B_j = b_j I_d, \quad (2.18)$$

donde  $a_i$  y  $b_j$  son constantes positivas.

### El modelo BEKK

Esta familia de modelos tiene la ventaja de que su construcción asegura que la matriz  $\Sigma_t$  sea definida positiva. El nombre de este modelo se obtiene de las iniciales de los nombres de sus autores Baba, Engle, Kroner y Kraft.

**Definición 2.24.** Diremos que el proceso  $(X_t)_{t \in \mathbb{Z}}$  es un proceso BEKK si es un proceso con la estructura general de la Definición 2.17, y la matriz de covarianzas condicional  $\Sigma_t$ , satisface que para toda  $t \in \mathbb{Z}$

$$\Sigma_t = A_0 + \sum_{i=1}^p A_i' X_{t-i} X_{t-i}' A_i + \sum_{j=1}^q B_j' \Sigma_{t-j} B_j, \quad (2.19)$$

donde las matrices  $A_i$  para  $i = 1, \dots, p$  y  $B_j$  para  $j = 1, \dots, q$  están en  $\mathbb{R}^{d \times d}$ , y  $A_0$  es simétrica y positiva definida.

**Proposición 2.4.** En el modelo BEKK, la matriz de covarianzas condicional  $\Sigma_t$  es positiva definida c.s. para toda  $t \in \mathbb{Z}$ .

*Demostración.* Por simplicidad, consideremos el modelo de primer orden. Para un vector  $v \neq 0$  en  $\mathbb{R}^d$ , tenemos

$$v' \Sigma_t v = v' A_0 v + (v' A_1' X_{t-1})^2 + (B_1 v)' \Sigma_{t-1} (B_1 v) > 0,$$

pues el primer término es estrictamente positivo y el segundo y tercero son no negativos.  $\square$

**Observación 2.6.** Existe una definición más general de la clase BEKK, que incluye todos los modelos DVEC. En esta definición tenemos

$$\Sigma_t = A_0 A_0' + \sum_{k=1}^K \sum_{i=1}^p A_{k,i}' X_{t-i} X_{t-i}' A_{k,i} + \sum_{k=1}^K \sum_{j=1}^q B_{k,j}' \Sigma_{t-j} B_{k,j},$$

donde  $d(d+1)/2 > K \geq 1$  y la elección de  $K$  determina la riqueza del modelo.

Sin embargo, este modelo tiende a ser muy complejo para aplicaciones prácticas, aún cuando  $K = 1$ , y es por eso que sólo se utiliza de manera teórica.

En la Tabla 2.1 están resumidas las fórmulas generales para el número de parámetros de cada modelo multivariado que hemos visto. De esta tabla podemos deducir que hay algunos modelos que no nos serán muy útiles (modelo VEC); otros sólo los usaremos cuando tengamos dimensiones pequeñas (modelos BEKK y DVEC); y unos más los usaremos de una manera más práctica, pero cuidando la dimensión, porque el número de parámetros crece rápidamente.

Modelo	Número Parámetros	2	5	10
VEC	$d(d+1)(1+(p+q)d(d+1)/2)/2$	21	465	6105
BEKK	$d(d+1)/2 + d^2(p+q)$	11	65	255
DVEC (2.16)	$d(d+1)(1+p+q)/2$	9	45	165
DCC	$d(d+1)/2 + (d+1)(p+q)$	9	27	77
CCC	$d(d+1)/2 + d(p+q)$	7	25	75
DVEC (2.17)	$d(d+1)/2 + d(p+q)$	7	25	75
DVEC (2.18)	$d(d+1)/2 + (p+q)$	5	17	57

**Tabla 2.1:** Parámetros de varios modelos GARCH Multivariados. La segunda columna da la fórmula general, y las siguientes son el número de dimensiones 2, 5 y 10 cuando  $p = 1 = q$ .

#### 2.5.4. Ajuste de modelos GARCH Multivariados

En la sección anterior vimos el ajuste de algunos modelos en etapas, y hay que enfatizar que cuando estamos trabajando en dimensiones muy altas, ésta puede ser la única estrategia viable. No obstante, si manejamos dimensiones más pequeñas (menores a 10), podemos intentar el ajuste de los modelos GARCH Multivariados en un sólo paso con una función de verosimilitud con respecto a todos los parámetros.

El método para construir una función de verosimilitud para cualquier modelo GARCH Multivariado  $X_t = \Sigma_t^{1/2} Z_t$  es análogo al caso univariado; por simplicidad lo haremos con el modelo de primer orden ( $p = 1 = q$ ), y supondremos que nuestros datos  $X_0, X_1, \dots, X_n$  están indexados. La máxima verosimilitud condicional está basada en la densidad conjunta de  $X_1, \dots, X_n$  dado  $X_0$ , y un valor inicial  $\Sigma_0$  para la matriz de covarianzas condicional. Esta

densidad conjunta condicional es

$$f_{X_1, \dots, X_n | X_0, \Sigma_0}(x_1, \dots, x_n | x_0, \Sigma_0) = \prod_{t=1}^n f_{X_t | X_{t-1}, \dots, X_0, \Sigma_0}(x_t | x_{t-1}, \dots, x_0, \Sigma_0).$$

Si denotamos la densidad de las innovaciones multivariadas  $Z_t$  por  $g(z)$ , entonces tenemos

$$f_{X_t | X_{t-1}, \dots, X_0, \Sigma_0}(x_t | x_{t-1}, \dots, x_0, \Sigma_0) = |\Sigma_t|^{-1/2} g(\Sigma_t^{-1/2} x_t),$$

donde  $\Sigma_t$  es una matriz que está en función de  $x_{t-1}, \dots, x_0$  y  $\Sigma_0$ . Usualmente, se supone que  $g(z)$  pertenece a la familia esférica, entonces  $g(z) = h(z'z)$  para alguna función  $h$ , y esto nos da una función de verosimilitud condicional de la forma

$$L(\theta; X_1, \dots, X_n) = \prod_{t=1}^n |\Sigma_t|^{-1/2} h(X_t' \Sigma_t^{-1} X_t),$$

donde todos los parámetros que aparecen en la ecuación de la volatilidad y la distribución de la innovación se reúnen en  $\theta$ . Además necesitamos dar un valor a  $\Sigma_0$ , que usualmente se escoge por métodos de maximización.

Para corroborar que nuestro modelo está bien, podemos revisar los residuales que se calculan a partir de  $\hat{Z}_t = \hat{\Sigma}_t^{-1/2} X_t$ , los cuales deben comportarse como una realización de un proceso SWN(0,  $I_d$ ).

### 2.5.5. PC-GARCH y O-GARCH

Supongamos que  $X_t = (X_{1,t}, \dots, X_{m,t})'$  es una serie de tiempo  $m$  dimensional con  $t = 1, \dots, T$ . Sea  $\chi$  una matriz de  $T \times m$  tal que sus filas están dadas por  $X_1', \dots, X_T'$ , y denotemos a las columnas de  $\chi$  por  $\chi_1, \dots, \chi_m$ . Ahora, si estandarizamos la matriz  $\chi$ , tendremos la matriz  $Y$  tal que cada una de sus columnas tendrá media cero y varianza 1, es decir, la  $i$ -ésima columna  $Y_i = (\chi_i - \mu_i)/\sigma_i$  donde  $\mu_i$  y  $\sigma_i$  son la media y la desviación estándar, respectivamente, de  $\chi_i$  para  $i = 1, \dots, m$ .

Sea  $\Sigma_Y$  la matriz de covarianzas de  $Y$ , si realizamos el análisis de componentes principales sobre  $Y$ , tendremos

$$P = Y\Gamma, \quad (2.20)$$

donde  $P$  denota la matriz de  $T \times m$  tal que sus columnas son las componentes principales; y  $\Gamma$  denota a los eigenvectores de la matriz de covarianzas  $\Sigma_Y$ . La matriz  $\Gamma$  estará ordenada de tal forma que la primer columna será el eigenvector correspondiente al eigenvalor más grande de la matriz de covarianzas  $\Sigma_Y$ ; la segunda columna será el eigenvector correspondiente al segundo eigenvalor más grande de  $\Sigma_Y$ ; y así sucesivamente.

Como  $\Gamma$  es una matriz ortogonal, es decir  $\Gamma' = \Gamma^{-1}$ , de (2.20) podemos deducir que

$$Y = P\Gamma'. \quad (2.21)$$

Al sacar las varianzas de (2.21), tenemos que

$$\Sigma_t = \Gamma\Delta_t\Gamma', \quad (2.22)$$

donde  $\Sigma_t$  es la matriz de covarianzas condicional de  $m \times m$  a tiempo  $t$ ; y  $\Delta_t$  es la matriz diagonal de las varianzas de las componentes principales de  $m \times m$  a tiempo  $t$ .

Las componentes principales sólo son no correlacionadas incondicionalmente, por lo que la matriz de covarianzas  $\Delta_t$  no será necesariamente diagonal, es por esto que suponemos que los elementos fuera de la diagonal de la matriz  $\Delta_t$  son cero, es decir, suponemos que las correlaciones condicionales son cero.

El modelo (2.22) es llamado PC-GARCH (GARCH de Componentes Principales) cuando la matriz diagonal de las varianzas de las componentes principales  $\Delta_t$  es estimada usando un modelo GARCH.

El principal motivo por el cual utilizamos el análisis de componentes principales es para reducir la dimensión, pues esperamos que un subgrupo de las componentes principales pueda explicar la mayor parte de la variabilidad presente en la matriz de covarianzas.

Es por eso que supondremos que las primeras  $r$  componentes explican la mayor parte de la variabilidad del sistema original, entonces podemos expresar a  $Y$  por

$$Y = P_r \Gamma_r' + \varepsilon, \quad (2.23)$$

donde  $P_r$  es la matriz con las primeras  $r$  componentes principales;  $\Gamma_r$  es la matriz con los  $r$  eigenvectores correspondientes a los  $r$  eigenvalores más grandes de  $\Sigma_Y$ ; y  $\varepsilon$  es el error que se produce al sólo usar las primeras  $r$  componentes. Usualmente en la práctica se ignora el término de error.

Al sacar las varianzas de (2.23), tenemos que

$$\Sigma_t = \Gamma_r \Delta_t \Gamma_r', \quad (2.24)$$

donde  $\Sigma_t$  es la matriz de covarianzas condicional de  $m \times m$  a tiempo  $t$ ; y  $\Delta_t$  es la matriz diagonal de las varianzas de las componentes principales de  $r \times r$  a tiempo  $t$ .

El modelo (2.24) es llamado O-GARCH (GARCH Ortogonal) cuando la matriz diagonal de las varianzas de las componentes principales  $\Delta_t$  es estimada usando un modelo GARCH. Sin embargo, pueden usarse otros métodos para estimar la matriz  $\Delta_t$ , como por ejemplo el EWMA. Al modelo que usa el EWMA para estimar la matriz  $\Delta_t$ , lo llamaremos el modelo O-EWMA.

El modelo O-GARCH requiere estimar  $r$  modelos GARCH por separado, uno para cada componente principal. Como  $\Delta_t$  siempre será positiva definida, la matriz del O-GARCH  $\Sigma_t$  siempre será positiva semidefinida. Para ver esto, escribamos

$$x' \Sigma_t x = x' \Gamma_r \Delta_t \Gamma_r' x = y' \Delta_t y, \quad (2.25)$$

donde  $y = \Gamma_r' x$ . Como  $y$  puede ser cero para alguna  $x \neq 0$ , entonces  $x' \Sigma_t x$  no será estrictamente positiva definida, pero sí positiva semidefinida.

Este método es muy eficiente, computacionalmente hablando, pues al realizar el análisis de componentes principales se reduce mucho la dimensión del problema, y basta con algunas varianzas GARCH para encontrar el resto de las varianzas y covarianzas del sistema.

Además presenta otra ventaja, pues permite la estimación por etapas. En primer lugar, calculamos la descomposición espectral de la matriz de

covarianzas de la muestra  $S$ , lo que nos da un estimador  $G$  de la matriz de los eigenvalores  $\Gamma$ . A continuación, rotamos los datos originales para obtener las componentes principales. Y para finalizar ajustamos un modelo GARCH univariado a cada componente principal y los residuales de nuestros modelos GARCH deberán comportarse como un  $\text{SWN}(0, I_d)$ .

# Capítulo 3

## Aplicación del modelo O-GARCH

En este capítulo mostraremos las ventajas que tiene el método O-GARCH sobre otros tipos de modelos GARCH Multivariados, al aplicarlo a datos financieros altamente correlacionados. Estos datos serán los precios de los *forwards* sobre el WTI y los precios de los *forwards* sobre el *Brent Blend*.

### 3.1. *Forwards* sobre el *West Texas Intermediate*

El *West Texas Intermediate* (WTI), también conocido como *Texas Light Sweet*, es una clase de petróleo bruto que se extrae de los campos occidentales del estado de Texas, Estados Unidos. Se dice que es “ligero” por su baja densidad, y “dulce” por su bajo contenido de azufre, siendo de los petróleos con menor cantidad de este elemento. Es un crudo de muy alta calidad que se utiliza para producir gasolina, destilados y otros derivados del petróleo. La mayor parte del tiempo, el WTI se refina directamente en Estados Unidos cerca de la Costa del Golfo. Su producción diaria es de 300 000 barriles, cuando en el mundo se producen un total de 85 millones de barriles al día.

Los *forwards* sobre el WTI se utilizan como precio de referencia internacional para fijar el valor de otros petróleos brutos como materias primas en los mercados de intercambio, principalmente en Norteamérica. Los primeros

*forwards* sobre el petróleo crudo fueron hechos en 1983, por la *Chicago Board of Trade* (CBOT) y por la *New York Mercantile Exchange* (NYMEX), tratando ambas de sacar ventaja de las regulaciones que el gobierno hacía sobre el petróleo crudo. CBOT tuvo problemas al principio con sus *forwards*, y la mayoría acabó yéndose con NYMEX.

El petróleo crudo se convirtió en el producto más comercializado a nivel mundial, mientras que los contratos *forwards* de NYMEX se volvieron la forma más líquida de intercambiar el petróleo crudo alrededor del mundo, al mismo tiempo que se posicionó como el contrato *forward* con mayor volumen de negociación en el mundo, llegando a comerciar alrededor de 850 000 contratos *forwards* diariamente. Además de los *forwards*, también se ofrecen otro tipo de oportunidades a través de opciones.

Los contratos se hacen sobre 1000 barriles de petróleo, y al vencimiento se puede exigir la entrega física o la compensación monetaria. El punto de entrega se sitúa en Cushing, Oklahoma (Estados Unidos), que es el punto más importante de la red de oleoductos estadounidense, y por la tanto un punto accesible para los mercados internacionales. Los contratos contemplan la entrega de varias calidades de crudos nacionales y extranjeros.

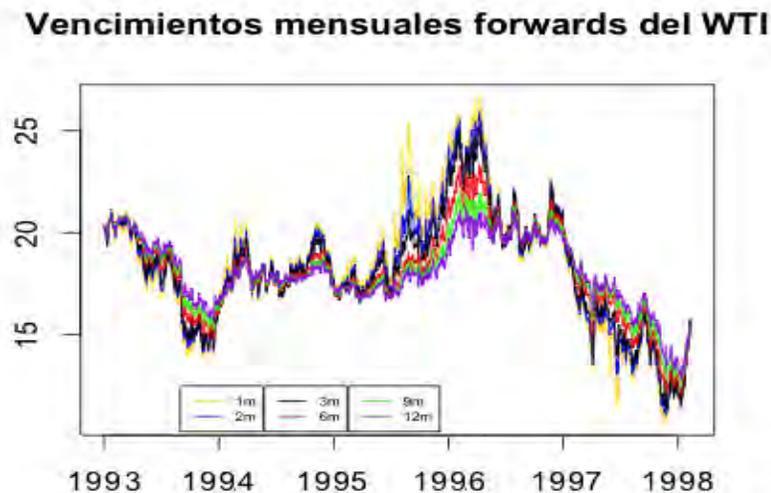
### 3.1.1. Aplicación

En Alexander (2001), la autora trabaja con los vencimientos mensuales, entre 1 mes y 12 meses, de los *forwards* del petróleo crudo del WTI; sus datos fueron recabados diariamente del 04/02/1993 al 24/03/1999 y pueden ser encontrados en su página personal <http://www.carolalexander.org/discussionpapers.php>.

En este artículo, se muestra la aplicación del método del O-GARCH a los datos del WTI. En primer lugar, replicaremos este estudio y mostraremos los resultados obtenidos y, posteriormente, haremos un estudio comparativo con datos actualizados con la misma ventana temporal, pero que fueron recabados del 10/01/2008 hasta el 26/02/2014; y extraídos del sitio <http://www.quandl.com/>.

**Datos del 04/02/1993 al 24/03/1999**

En la Figura 3.1, podemos observar los precios históricos, de algunos vencimientos, de los forwards sobre el WTI.



**Figura 3.1:** Vencimientos a 1, 2, 3, 6, 9 y 12 meses de los forwards del petróleo crudo del WTI.

Previo al análisis, observemos la correlación de los datos. En la Tabla 1.3 (Capítulo 1), podemos notar que los datos están fuertemente correlacionados, por lo que es de esperar un buen desempeño del modelo.

En cambio, en la Tabla 1.4 (Capítulo 1), se observa que con tan sólo las primeras 3 componentes se explica más del 99.8% de la varianza total. De hecho, con la primer componente se puede explicar casi el 95% de la varianza del sistema, y con dos el 99%; pero al ver la Gráfica de Codo de la Figura 1.4 (Capítulo 1), nos podemos dar cuenta que basta con usar las primeras 2 componentes principales, porque después de la tercer componente, los datos ya son muy parecidos en magnitud y ahí es donde encontramos esa especie de “codo” que nos interesa en este tipo de gráficas.

La gran ventaja de usar el método del O-GARCH, es que todas las volatilidades y correlaciones del sistema las podremos derivar de tan sólo 2 varianzas GARCH con simples transformaciones en la matriz diagonal de las varianzas de las componentes principales, de esta manera resulta fácil estimar las 78 volatilidades y correlaciones.

En la Tabla 3.1, se muestra el ajuste de los modelos GARCH(1,1) a las primeras 2 componentes principales.

	P1	P2
$\mu$	-3.077292e-18	-1.315937e-19
$\alpha_0$	1.891144e-05	1.492057e-05
$\alpha_1$	0.04615184	0.3115531
$\beta_1$	0.9479762	0.5491947

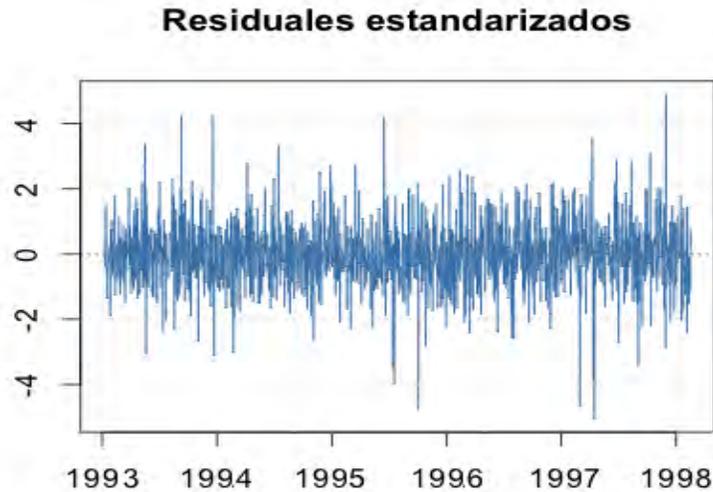
**Tabla 3.1:** Modelos GARCH(1,1) de las primeras dos componentes principales.

Ahora examinaremos si el ajuste del GARCH(1,1) a nuestras primeras dos componentes principales es correcto; cabe resaltar que este análisis ya no es parte del artículo de la Dra. Alexander.

Primero veamos el ajuste del GARCH(1,1) a la primer componente principal. En la Figura 3.2 podemos observar la serie de tiempo de los residuales que al parecer no se comporta como una normal, sin embargo necesitamos de otro gráfico para rechazar esta hipótesis.

En la Figura 3.3, tenemos el Gráfico *QQ* de los residuales de la primer componente contra una distribución normal estándar, y por el comportamiento de ambas colas podemos decir que las innovaciones no se comportan como una normal estándar, de hecho las colas tienen un mayor peso de lo que tendría una normal, por lo que necesitamos una distribución de las innovaciones que cubra esta necesidad. Entonces, ajustaremos un GARCH(1,1) con innovaciones  $t$ .

Como podemos ver en la Figura 3.4, el ajuste es muchísimo mejor que cuando supusimos la normalidad de las innovaciones, incluso, salvo algunos



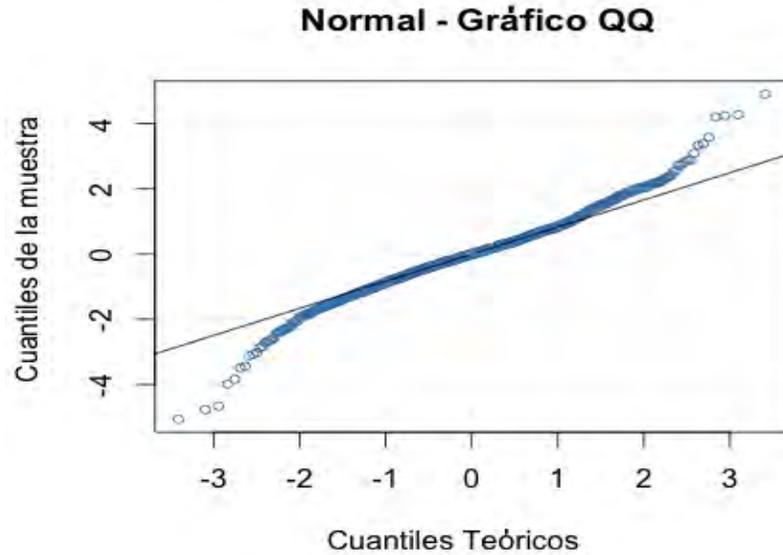
**Figura 3.2:** Residuales estandarizados de la primer componente principal.

puntos, los residuales se comportan como una  $t$  de Student con 4.92 grados de libertad.

Al hacer el análisis del ajuste del GARCH(1,1) a la segunda componente, notamos que los residuales se ajustan mejor si suponemos que las innovaciones se distribuyen  $t$  de Student. En la Figura 3.5 podemos observar los Gráficos  $QQ$  de ambos ajustes, cuando suponemos normalidad en las innovaciones y cuando suponemos que se distribuyen  $t$  de Student.

La Tabla 3.2 muestra los parámetros del ajuste del GARCH(1,1) con innovaciones  $t$  a las primeras 2 componentes principales. Haciendo énfasis en que R estima los grados de libertad de la distribución  $t$  de Student.

Ahora es momento de usar el método O-GARCH para estimar la matriz de varianzas y covarianzas del sistema. En la Figura 3.6 mostramos algunas de las volatilidades que son generadas al usar el método O-GARCH comparadas con las estimaciones directas de las volatilidades usando un GARCH(1,1). Como podemos observar, ambas volatilidades son muy parecidas.



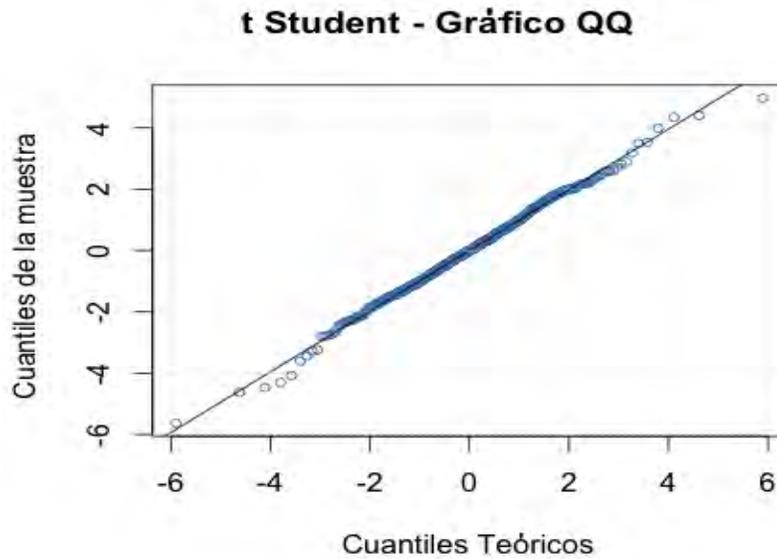
**Figura 3.3:** Gráfico  $QQ$  de los residuales estandarizados de la primer componente al ajustar un  $GARCH(1,1)$  contra una distribución normal estándar.

	P1	P2
$\mu$	-3.077292e-18	-1.315937e-19
$\alpha_0$	4.966836e-06	1.551768e-05
$\alpha_1$	0.02660327	0.2769823
$\beta_1$	0.9733750	0.5827984
$\nu$	4.920403	4.332639

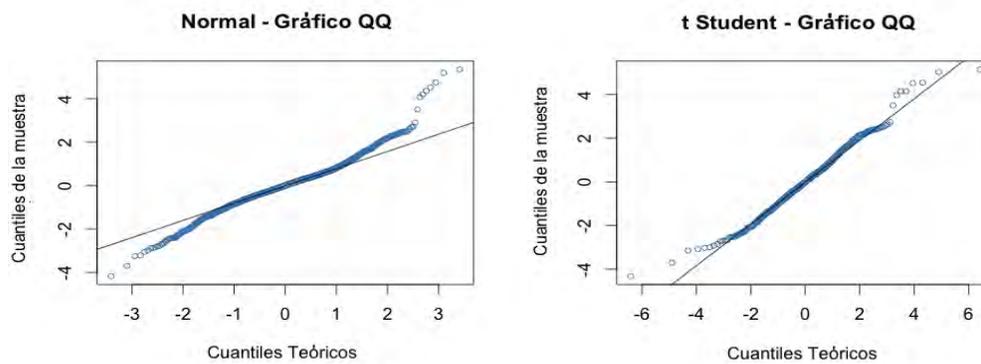
**Tabla 3.2:** Modelos  $GARCH(1,1)$  con innovaciones  $t$  de las primeras dos componentes principales, donde  $\nu$  representa los grados de libertad de la distribución  $t$ .

Mientras tanto, en la Figura 3.7, se muestra una comparación del método O-EWMA con el método O-GARCH. Nos podemos dar cuenta que las volatilidades EWMA son muy parecidas a las del O-GARCH. Entonces, ¿por qué molestarse en usar modelos GARCH?

La razón es que las correlaciones del GARCH reflejan de una manera



**Figura 3.4:** Gráfico *QQ* de los residuos estandarizados de la primer componente al ajustar un GARCH(1,1) con innovaciones *t* contra una distribución *t* de Student con 4.92 grados de libertad.

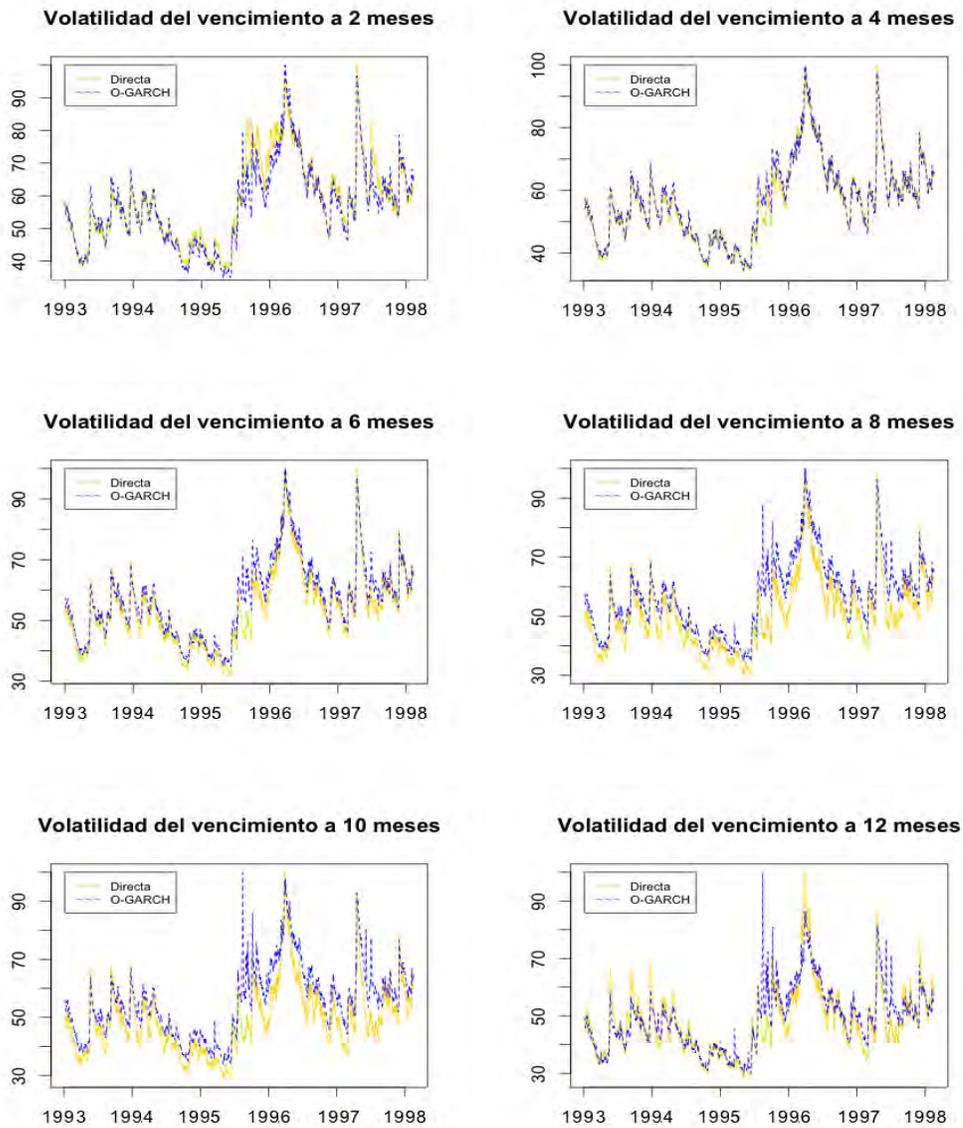


**Figura 3.5:** Del lado izquierdo podemos observar el Gráfico *QQ* de los residuos estandarizados de la segunda componente al ajustar un GARCH(1,1) contra una distribución normal estándar; mientras que del lado derecho podemos observar el Gráfico *QQ* de los residuos estandarizados de la segunda componente al ajustar un GARCH(1,1) con innovaciones *t* contra una distribución *t* de Student con 4.33 grados de libertad.

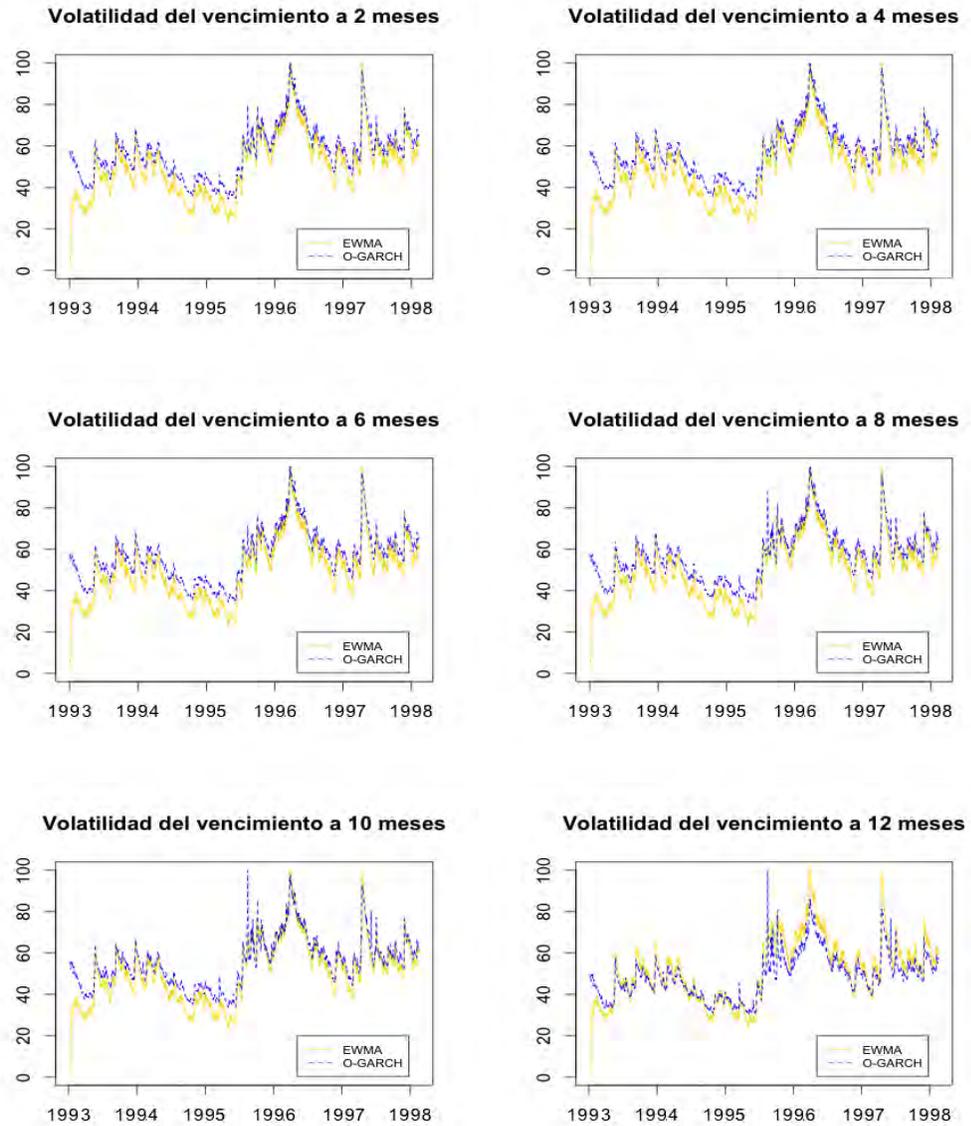
más realista el comportamiento del mercado. Esto se debe al valor que se escoge para la constante de suavidad en el modelo EWMA, ya que si uno reduce el valor de la constante, al mismo tiempo reduce la persistencia de las volatilidades y de las correlaciones, y encontrar el valor apropiado para esta constante es muy difícil.

Cuando parametrizamos modelos GARCH Multivariados se suelen comparar las volatilidades GARCH, que se estiman con el modelo multivariado GARCH, con aquellas obtenidas de estimar directamente con un modelo univariado GARCH. De manera similar, cuando se evalúa un modelo O-GARCH, uno puede comparar las volatilidades y las correlaciones estimadas con aquellas obtenidas de otros modelos, como el modelo O-EWMA o algún otro modelo GARCH Multivariado. Sin embargo, surgen problemas como el de saber con qué constante de suavidad se trabajará, si usamos el modelo O-EWMA; o si el sistema es muy grande, seguramente tendremos problemas de convergencia, entonces no podremos confiar mucho en las estimaciones que obtengamos de otros modelos GARCH Multivariados.

Para saber si nuestro modelo ha sido correctamente ajustado, podemos revisar los residuales de los modelos GARCH que se le ajustaron a las componentes principales, si estos se comportan como un ruido blanco estricto, entonces podemos decir que nuestro modelo ha sido correctamente ajustado. En la Figura 3.8, se indica que los ACF y PACF de nuestras componentes principales se comportan como un ruido blanco estricto, por lo tanto nuestro modelo está bien ajustado.



**Figura 3.6:** Comparaciones de las volatilidades de los vencimientos 2, 4, 6, 8, 10 y 12 meses de los forwards del petróleo crudo del WTI, calculados directamente y usando el método O-GARCH.



**Figura 3.7:** Comparaciones de las volatilidades de los vencimientos 2, 4, 6, 8, 10 y 12 meses de los forwards del petróleo crudo del WTI, usando el método O-EWMA y el método O-GARCH.

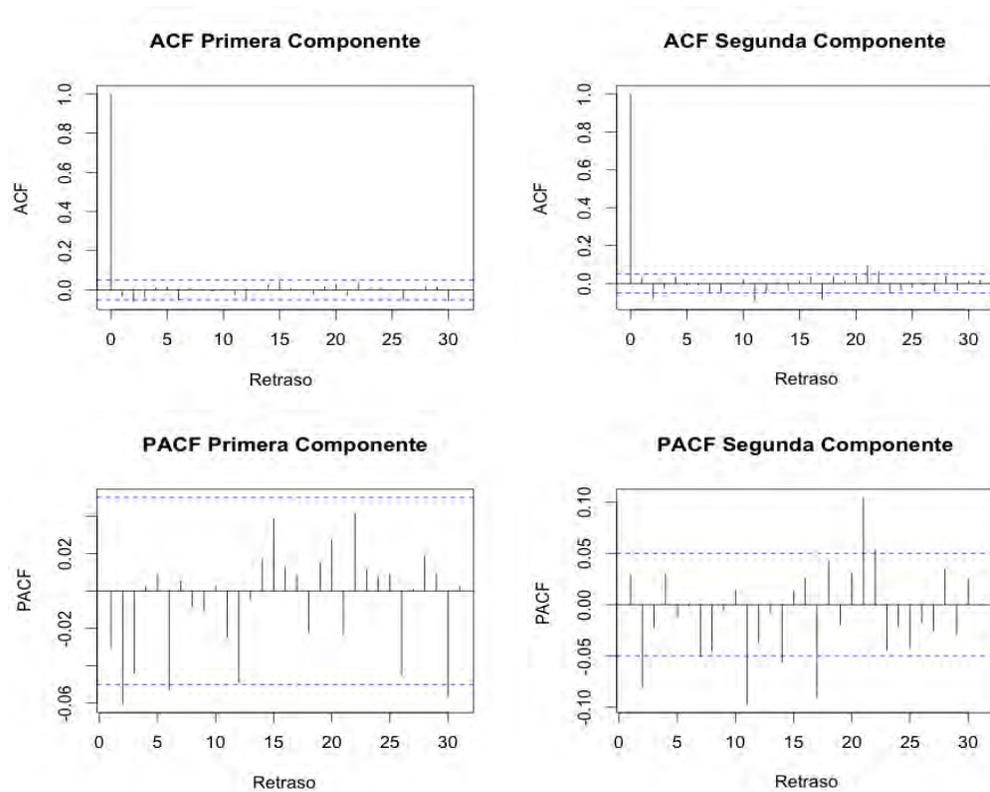


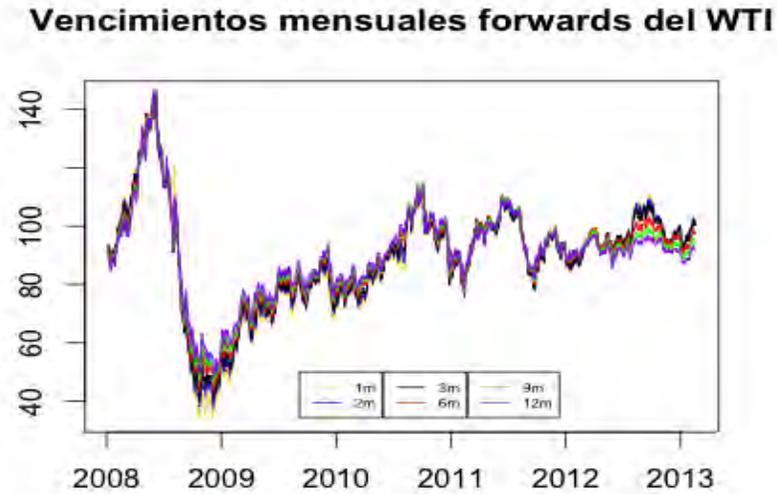
Figura 3.8: ACF y PACF de los residuales de las primeras dos componentes principales.

### Datos del 10/01/2008 al 26/02/2014

En la Figura 3.9, podemos observar los precios históricos, de algunos vencimientos, de los *forwards* sobre el WTI. Cabe resaltar, que los precios están dados en dólares.

Para empezar nuestro análisis, primero veamos que tan correlacionadas están nuestras variables. En la Tabla 3.3 podemos ver las correlaciones de los log-rendimientos de los *forwards*. Hay que notar que las variables están altamente correlacionadas entre sí, por lo que es de esperar que el análisis de componentes principales se desempeñe muy bien, y por lo tanto el modelo O-GARCH también.

En la Tabla 3.4 indica que la primer componente principal, tan sólo expli-



**Figura 3.9:** Vencimientos a 1, 2, 3, 6, 9 y 12 meses de los forwards del petróleo crudo del WTI.

	1m	2m	3m	4m	5m	6m	7m	8m	9m	10m	11m	12m
1m	1											
2m	0.9488	1										
3m	0.9292	0.9952	1									
4m	0.9178	0.9890	0.9984	1								
5m	0.9093	0.9833	0.9955	0.9991	1							
6m	0.9018	0.9778	0.9920	0.9971	0.9993	1						
7m	0.8958	0.9734	0.9888	0.9949	0.9980	0.9995	1					
8m	0.8909	0.9694	0.9856	0.9925	0.9963	0.9985	0.9996	1				
9m	0.8860	0.9650	0.9821	0.9897	0.9941	0.9969	0.9987	0.9996	1			
10m	0.8815	0.9608	0.9787	0.9869	0.9918	0.9952	0.9974	0.9988	0.9997	1		
11m	0.8765	0.9562	0.9749	0.9837	0.9892	0.9930	0.9957	0.9976	0.9989	0.9997	1	
12m	0.8713	0.9515	0.9709	0.9804	0.9864	0.9907	0.9938	0.9961	0.9978	0.9990	0.9997	1

**Tabla 3.3:** Matriz de correlaciones de los log-rendimientos de los vencimientos mensuales de los *forwards* del WTI.

ca el 97 % de la varianza del sistema original y con dos componentes podemos explicar el 99 % de la varianza del sistema original. Las cantidades resultan sorprendentes, pero era lo que esperábamos dada la correlación entre las variables.

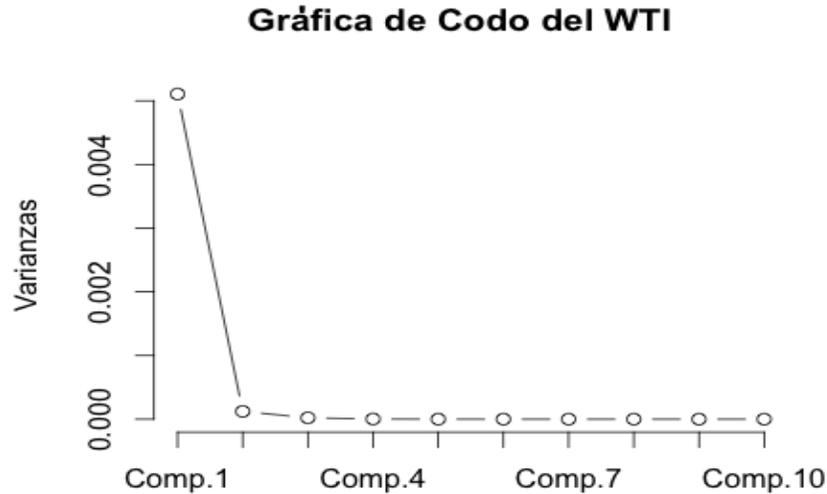
$\lambda_i$	Proporción Varianza	Proporción Acumulada
5.107185e-03	0.972062600	0.9720626
1.220307e-04	0.023226390	0.9952890
2.114397e-05	0.004024382	0.9993134
2.819381e-06	0.000536619	0.9998500
5.132895e-07	0.000097695	0.9999477
1.443557e-07	0.000027475	0.9999752
6.306341e-08	0.000012003	0.9999872
3.124494e-08	0.000005946	0.9999931
1.371660e-08	0.000002610	0.9999958
1.067380e-08	0.000002031	0.9999978
6.514745e-09	0.000001239	0.9999999
5.126965e-09	0.000000975	1

**Tabla 3.4:** Eigenvalores de las componentes principales de los log-rendimientos de los vencimientos mensuales de los *forwards* del WTI.

La Gráfica de Codo que se muestra en la Figura 3.10, aclara que basta con usar exclusivamente la primer componente principal para explicar la varianza del sistema, ya que después de ésta los datos son muy parecidos en magnitud y en la segunda componente encontramos esa especie de “codo” que nos interesa en esta gráfica. Además, sabemos de antemano que la primer componente explica el 97 % de la varianza del sistema.

Sin embargo, si sólo usáramos la primer componente principal, tendríamos que las correlaciones serían unitarias, así que toda la variación en la correlación proviene de los movimientos en la segunda componente principal. Esta segunda componente principal explica el 2.3 % del movimiento.

Ahora que ya decidimos que sólo usaremos las primeras dos componentes, ajustaremos un modelo GARCH(1,1) a cada component. En la Tabla 3.5 se muestra el ajuste de los modelos GARCH(1,1) para las primeras dos componentes principales de este sistema, podemos observar que la segunda componente principal tiene mucha reacción (dado que  $\alpha$  es cercana a 0.23), pero poca persistencia (dado que  $\beta$  es cercana a 0.76). En otras palabras estos movimientos son intensos, pero duran poco.



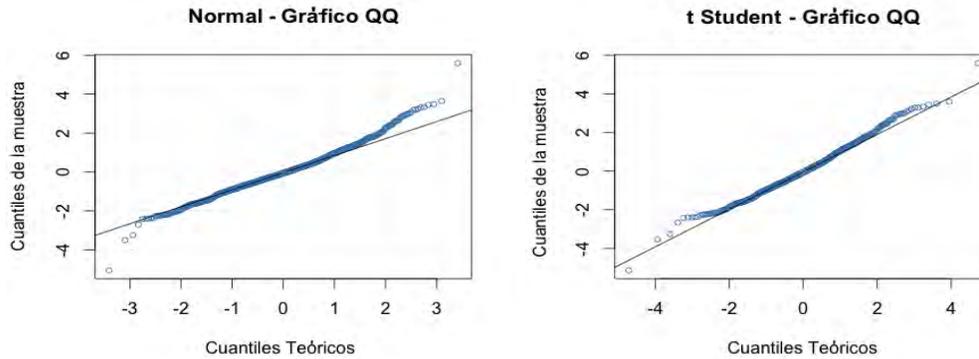
**Figura 3.10:** Gráfica de Codo de los log-rendimientos de los vencimientos mensuales de los forwards del WTI.

	P1	P2
$\mu$	2.697500e-18	-2.092320e-18
$\alpha_0$	2.089698e-05	1.665892e-06
$\alpha_1$	0.06148257	0.2296515
$\beta_1$	0.9349286	0.7578918

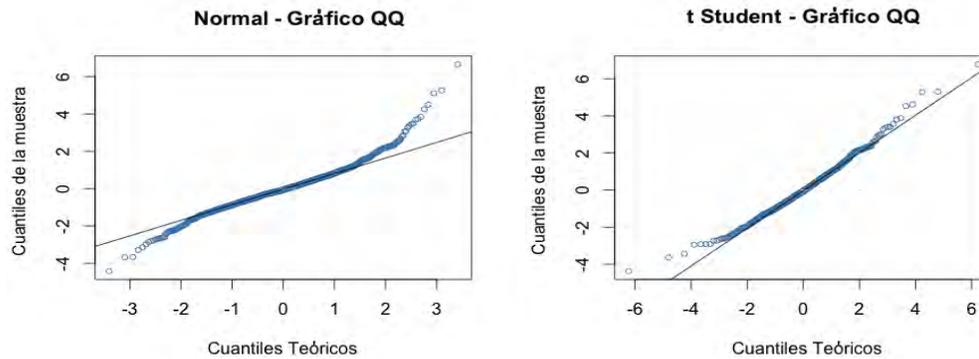
**Tabla 3.5:** Modelos GARCH(1,1) de las primeras dos componentes principales.

Analicemos, primero, el ajuste del GARCH(1,1) a la primer componente principal. En la Figura 3.11 se ven los Gráficos  $QQ$  de ajustar un GARCH(1,1) con innovaciones normales y otro de innovaciones  $t$  de Student. Es claro que el ajuste es mucho mejor cuando utilizamos las innovaciones  $t$  de Student.

Ahora, al hacer el análisis del ajuste del GARCH(1,1) de la segunda componente, queda claro que los residuales se ajustan mejor de una manera más sorprendente que con la primer componente, si suponemos que las innovaciones se distribuyen  $t$  de Student. En tanto que en la Figura 3.12 muestra los



**Figura 3.11:** Del lado izquierdo podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la primera componente al ajustar un GARCH(1,1) contra una distribución normal estándar; mientras que del lado derecho podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la primera componente al ajustar un GARCH(1,1) con innovaciones  $t$  contra una distribución  $t$  de Student con 7.75 grados de libertad.



**Figura 3.12:** Del lado izquierdo podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la segunda componente al ajustar un GARCH(1,1) contra una distribución normal estándar; mientras que del lado derecho podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la segunda componente al ajustar un GARCH(1,1) con innovaciones  $t$  contra una distribución  $t$  de Student con 4.52 grados de libertad.

Gráficos  $QQ$  de ambos ajustes, cuando suponemos normalidad en las innovaciones y cuando suponemos que se distribuyen  $t$  de Student.

En la Tabla 3.6 se muestran los parámetros del ajuste del GARCH(1,1) con innovaciones  $t$  a las primeras 2 componentes principales.

	P1	P2
$\mu$	2.697500e-18	-2.092320e-18
$\alpha_0$	1.531386e-05	1.128826e-06
$\alpha_1$	0.05560653	0.1801898
$\beta_1$	0.9419250	0.8147536
$\nu$	7.754387	4.526876

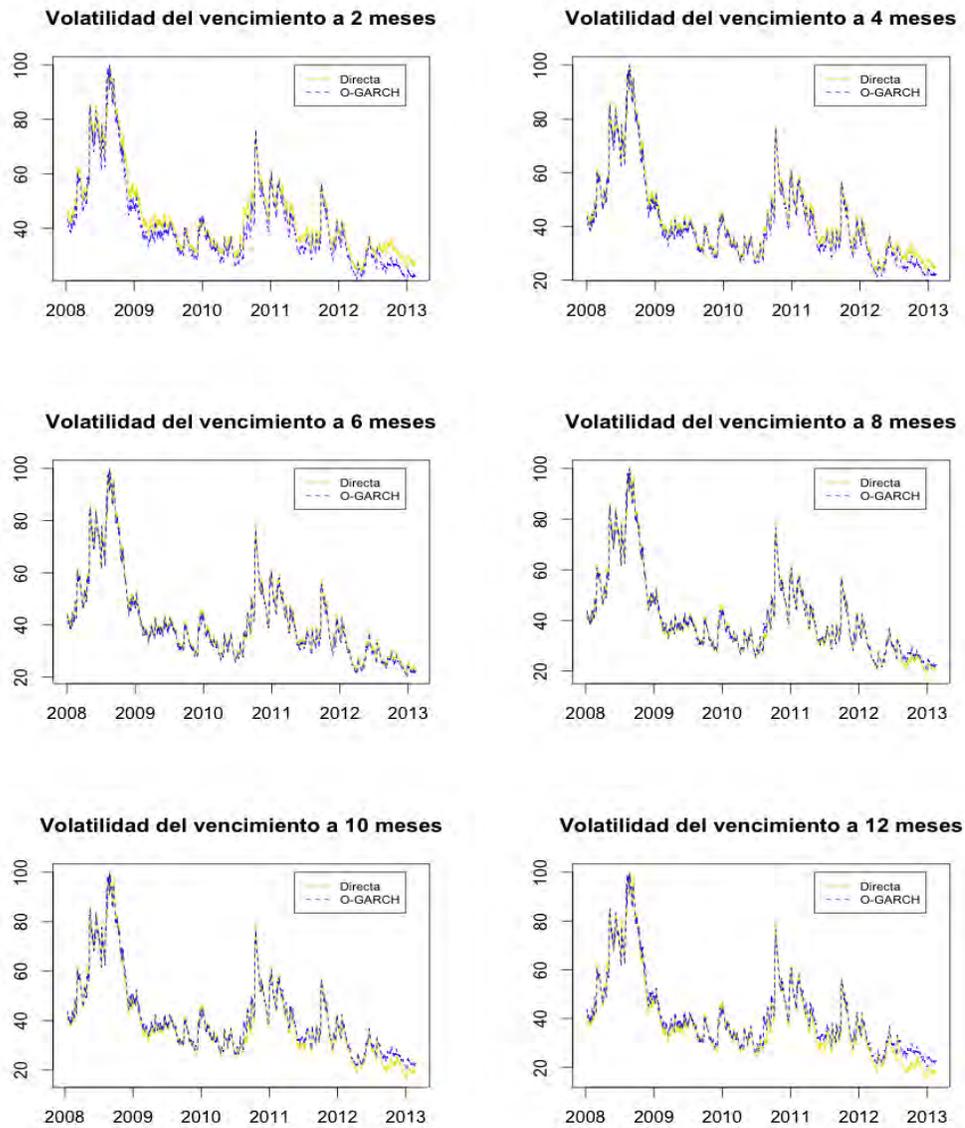
**Tabla 3.6:** Modelos GARCH(1,1) con innovaciones  $t$  de las primeras dos componentes principales, donde  $\nu$  representa los grados de libertad de la distribución  $t$ .

En la Figura 3.13 mostramos algunas de las volatilidades que son generadas al usar el método O-GARCH comparadas con las estimaciones directas de las volatilidades usando un GARCH(1,1).

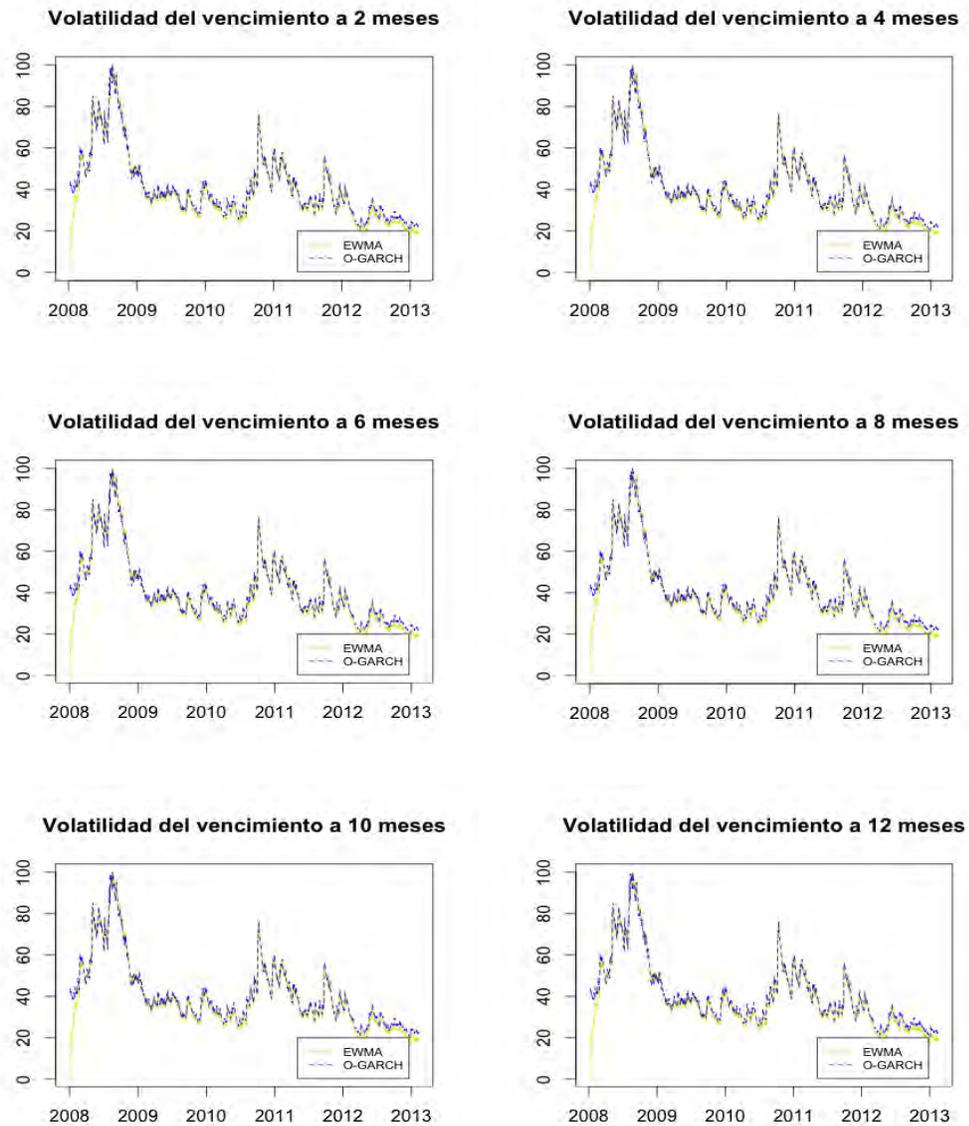
Mientras tanto, en la Figura 3.14, se muestra una comparación del método O-EWMA con el método O-GARCH. Se puede apreciar que ambas volatilidades son prácticamente idénticas. Por lo que podemos esperar que nuestra estimación usando el método O-GARCH sea buena.

Para corroborar que nuestro ajuste esté bien hecho, observamos los residuales de los ajustes de nuestras componentes principales, si estos se comportan como un ruido blanco estricto, entonces podemos asegurar que nuestro modelo está bien. En la Figura 3.15 podemos ver el ACF y PACF de las componentes principales, donde sus residuales se comportan como un ruido blanco estricto, podemos asegurar que el modelo es el correcto.

Al comparar los dos análisis, es decir, el que replicamos de Alexander (2001) y los de los datos actualizados, podemos ver que el precio del petróleo ha subido mucho desde 1999, y gran parte de esa alza se debe a los conflictos bélicos existentes en el mundo, de hecho actualmente hay 22 países en guerra. Esta alza no sólo se debe a las guerras, el petróleo es un producto fundamental para el ser humano y esta demanda hace también que los precios suban. Igualmente podemos notar que los parámetros de los modelos GARCH(1,1) son bastante parecidos, por lo que se puede esperar que los comportamientos sean parecidos.



**Figura 3.13:** Comparaciones de las volatilidades de los vencimientos 2, 4, 6, 8, 10 y 12 meses de los forwards del petróleo crudo del WTI, calculados directamente y usando el método O-GARCH.



**Figura 3.14:** Comparaciones de las volatilidades de los vencimientos 2, 4, 6, 8, 10 y 12 meses de los forwards del petróleo crudo del WTI, usando el método O-EWMA y el método O-GARCH.

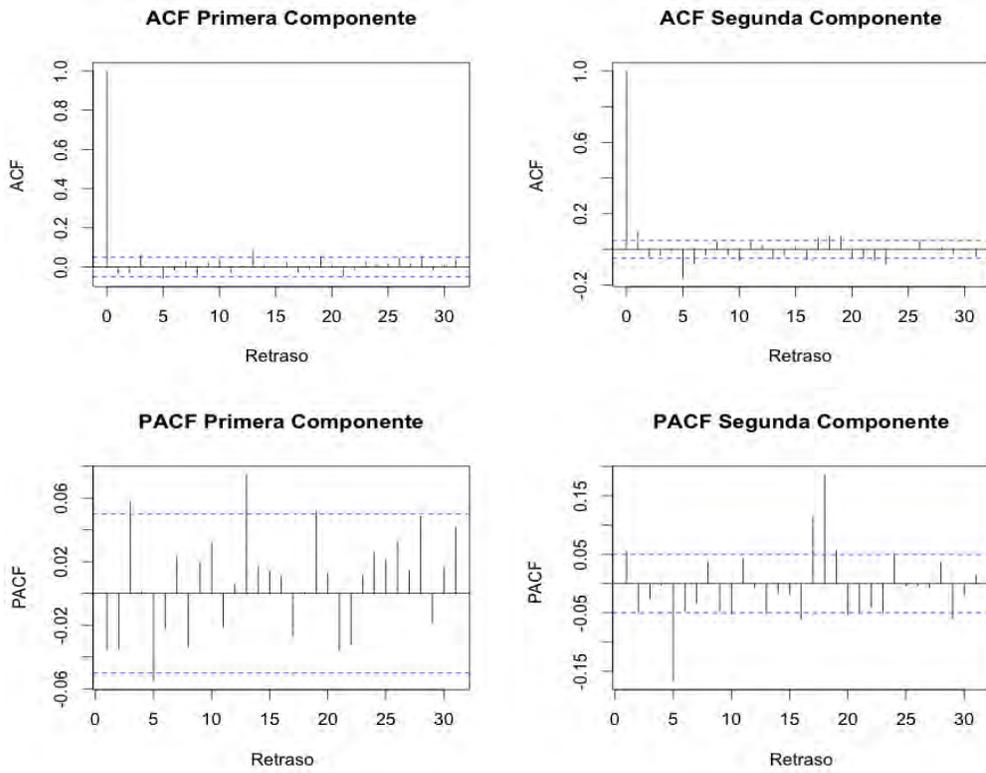


Figura 3.15: ACF y PACF de los residuales de las primeras dos componentes principales.

## 3.2. *Forwards* sobre el *Brent Blend*

Existe también otro punto de referencia muy importante para el crudo del petróleo, éste es el *Brent Blend*, que es el segundo más usado después del WTI, salvo que el *Brent Blend* es utilizado principalmente en Europa y África. Cabe resaltar que 65% de las variedades de crudo en el mercado mundial lo toman como precio de referencia.

El *Brent Blend* es una mezcla de 15 diferentes tipos de crudo provenientes de 15 diferentes campos de petróleo en el Mar del Norte. Su nombre se debe a un tipo de ganso, cuyo nombre en inglés es *Brent Goose*, que habita el Mar del Norte y que tiene la cabeza y el cuello del mismo color que el petróleo; este nombre fue dado en los años setenta por las compañías *ExxonMobil* y *Royal Dutch Shell* que explotaban el yacimiento.

Este crudo es de una calidad muy alta, debido a que es “ligero” y “dulce”, aunque no tanto como el WTI; y es utilizado principalmente para la elaboración de gasolina. Se suele refinar en los países de Europa Noroccidental, pero cuando los precios del mercado son lo suficientemente bajos para exportarlo, es procesado en las refinerías de la zona mediterránea y la costa este de Estados Unidos. Incluso el *Brent Blend* constituye más de la mitad del suministro mundial de petróleo crudo comercializado.

Los *forwards* sobre el *Brent Blend* comenzaron en junio de 1988 en Londres a través de la *Open Outcry International Petroleum Exchange*, pero desde el 2005, se efectúan de manera electrónica a través del ICE (*IntercontinentalExchange*). Actualmente el *Brent Blend* también cotiza en NYMEX. Además de los *forwards*, existen otro tipo de opciones sobre este crudo. Cada contrato *forward* se hace sobre 1000 barriles y al vencimiento se puede exigir la entrega física o la compensación monetaria. Los precios de los contratos están dados en dólares.

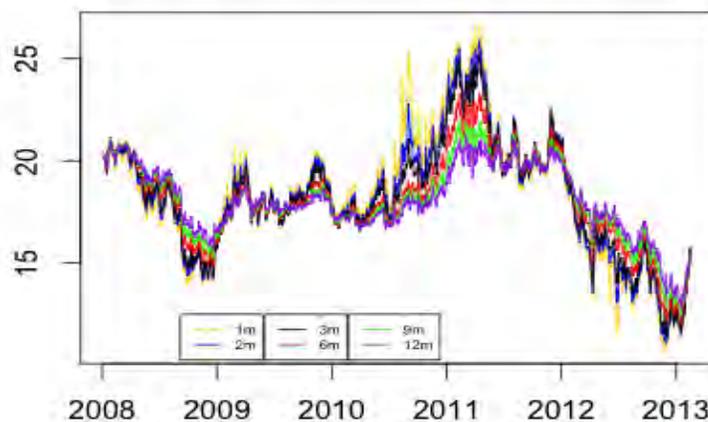
### 3.2.1. Aplicación

Ahora trabajaremos con los vencimientos mensuales, entre 1 mes y 12 meses, de los *forwards* del petróleo crudo *Brent Blend*; los datos fueron recabados diariamente del 28/03/2008 al 26/02/2014. Considerando la misma

ventana de tiempo que en el caso del WTI. La información fue tomada del mismo sitio que la de WTI.

En la Figura 3.16 podemos observar los precios históricos, de algunos vencimientos, de los forwards sobre el *Brent Blend*.

**Vencimientos mensuales forward BrentBlend**



**Figura 3.16:** Vencimientos a 1, 2, 3, 6, 9 y 12 meses de los forwards del petróleo crudo del *Brent Blend*.

En la Tabla 3.7, se muestran las correlaciones que tienen estas variables, las cuales son muy altas, por lo que esperamos que al realizar el análisis de componentes principales, con muy pocas componentes podamos explicar la mayor parte de la varianza del sistema, tal y como sucedió en el ejemplo anterior.

Como era de esperarse, al ver la Tabla 3.8, nos damos cuenta que la primer componente principal explica casi el 95% de la varianza del sistema original y con las primeras dos componentes principales explicamos el 99% de la varianza del sistema.

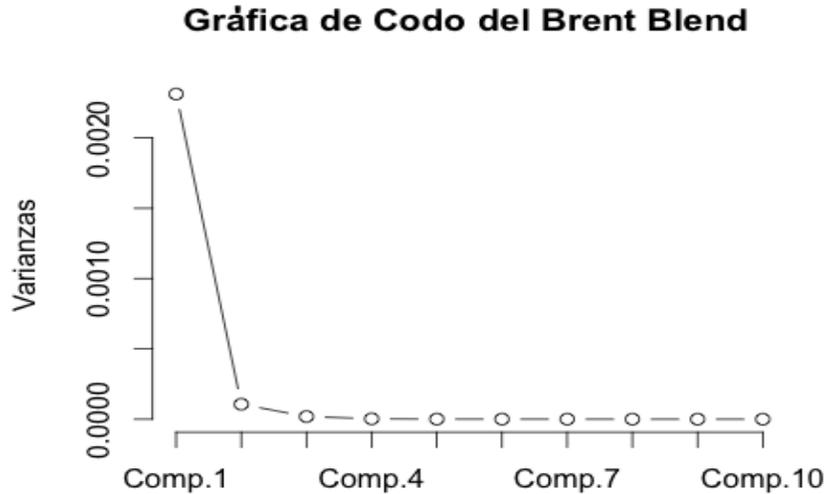
	1m	2m	3m	4m	5m	6m	7m	8m	9m	10m	11m	12m
1m	1											
2m	0.9500	1										
3m	0.9295	0.9936	1									
4m	0.9105	0.9827	0.9960	1								
5m	0.8928	0.9702	0.9884	0.9974	1							
6m	0.8781	0.9585	0.9801	0.9925	0.9983	1						
7m	0.8646	0.9468	0.9709	0.9860	0.9944	0.9984	1					
8m	0.8517	0.9349	0.9609	0.9783	0.9889	0.9949	0.9986	1				
9m	0.8382	0.9225	0.9503	0.9698	0.9825	0.9903	0.9957	0.9988	1			
10m	0.8250	0.9103	0.9395	0.9608	0.9752	0.9846	0.9915	0.9961	0.9989	1		
11m	0.8083	0.8941	0.9248	0.9476	0.9637	0.9747	0.9832	0.9893	0.9934	0.9960	1	
12m	0.7975	0.8841	0.9160	0.9402	0.9578	0.9701	0.9797	0.9870	0.9923	0.9961	0.9960	1

**Tabla 3.7:** Matriz de correlaciones de los log-rendimientos de los vencimientos mensuales de los *forwards* del *Brent Blend*.

$\lambda_i$	Proporción Varianza	Proporción Acumulada
2.312412e-03	0.947045400	0.9470454
1.061673e-04	0.043480690	0.9905261
1.868485e-05	0.007652357	0.9981784
2.679683e-06	0.001097461	0.9992759
6.722541e-07	0.000275320	0.9995512
4.583032e-07	0.000187697	0.9997389
2.948113e-07	0.000120739	0.9998597
1.547259e-07	0.000063367	0.9999230
8.352090e-08	0.000034205	0.9999572
4.799800e-08	0.000019657	0.9999769
3.503763e-08	0.000014349	0.9999912
2.140318e-08	0.000008765	1

**Tabla 3.8:** Eigenvalores de las componentes principales de los log-rendimientos de los vencimientos mensuales de los *forwards* del *Brent Blend*.

Al igual que en el ejemplo anterior, podemos restringirnos a sólo usar una componente principal, como lo sugiere la Gráfica de Codo de la Figura 3.17, sin embargo, decidimos quedarnos con las primeras dos componentes principales, porque al observar la Tabla 3.9, tenemos que la segunda componente principal tiene mucha reacción (dado que  $\alpha$  es cercana a 0.31), pero poca persistencia (dado que  $\beta$  es cercana a 0.55).



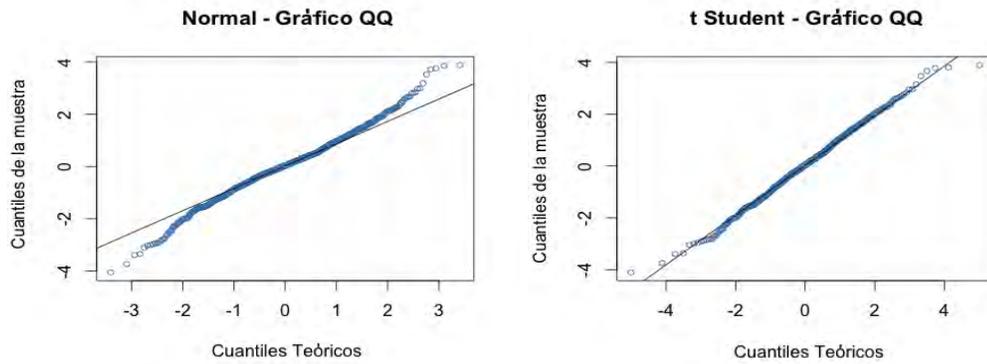
**Figura 3.17:** Gráfica de Codo de los log-rendimientos de los vencimientos mensuales de los forwards del *Brent Blend*.

	P1	P2
$\mu$	2.799777e-18	1.230609e-18
$\alpha_0$	2.978565e-05	5.764784e-07
$\alpha_1$	0.08014074	0.1079868
$\beta_1$	0.9156398	0.8788180

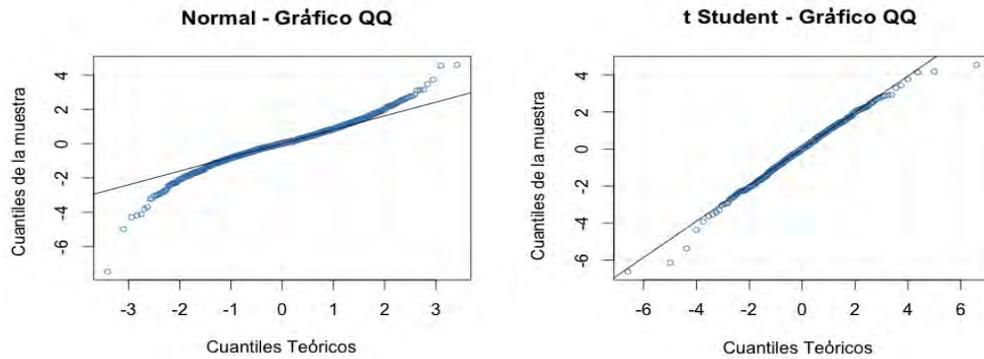
**Tabla 3.9:** Modelos GARCH(1,1) de las primeras dos componentes principales.

Ahora veamos si es correcto suponer que las innovaciones sean normales. Para la primer componente podemos observar en la Figura 3.18, el Gráfico *QQ* de los residuales del ajuste del GARCH(1,1) con innovaciones normales y con innovaciones *t* de Student. Es claro que es mejor suponer que las innovaciones sean una *t* de Student.

Para la segunda componente realizamos el mismo análisis, y como podemos observar en la Figura 3.19 es mucho mejor suponer que las innovaciones son *t* de Student.



**Figura 3.18:** Del lado izquierdo podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la primer componente al ajustar un GARCH(1,1) contra una distribución normal estándar; mientras que del lado derecho podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la primer componente al ajustar un GARCH(1,1) con innovaciones  $t$  contra una distribución  $t$  de Student con 6.69 grados de libertad.



**Figura 3.19:** Del lado izquierdo podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la segunda componente al ajustar un GARCH(1,1) contra una distribución normal estándar; mientras que del lado derecho podemos observar el Gráfico  $QQ$  de los residuales estandarizados de la segunda componente al ajustar un GARCH(1,1) con innovaciones  $t$  contra una distribución  $t$  de Student con 4.17 grados de libertad.

En la Tabla 3.10 se indican los nuevos parámetros de las 2 componentes principales que se obtienen al ajustar un modelo GARCH(1,1) con innovaciones  $t$  de Student.

	P1	P2
$\mu$	2.799777e-18	1.230609e-18
$\alpha_0$	2.353438e-05	1.876229e-06
$\alpha_1$	0.07560316	0.2287947
$\beta_1$	0.9234508	0.7407447
$\nu$	6.696350	4.171444

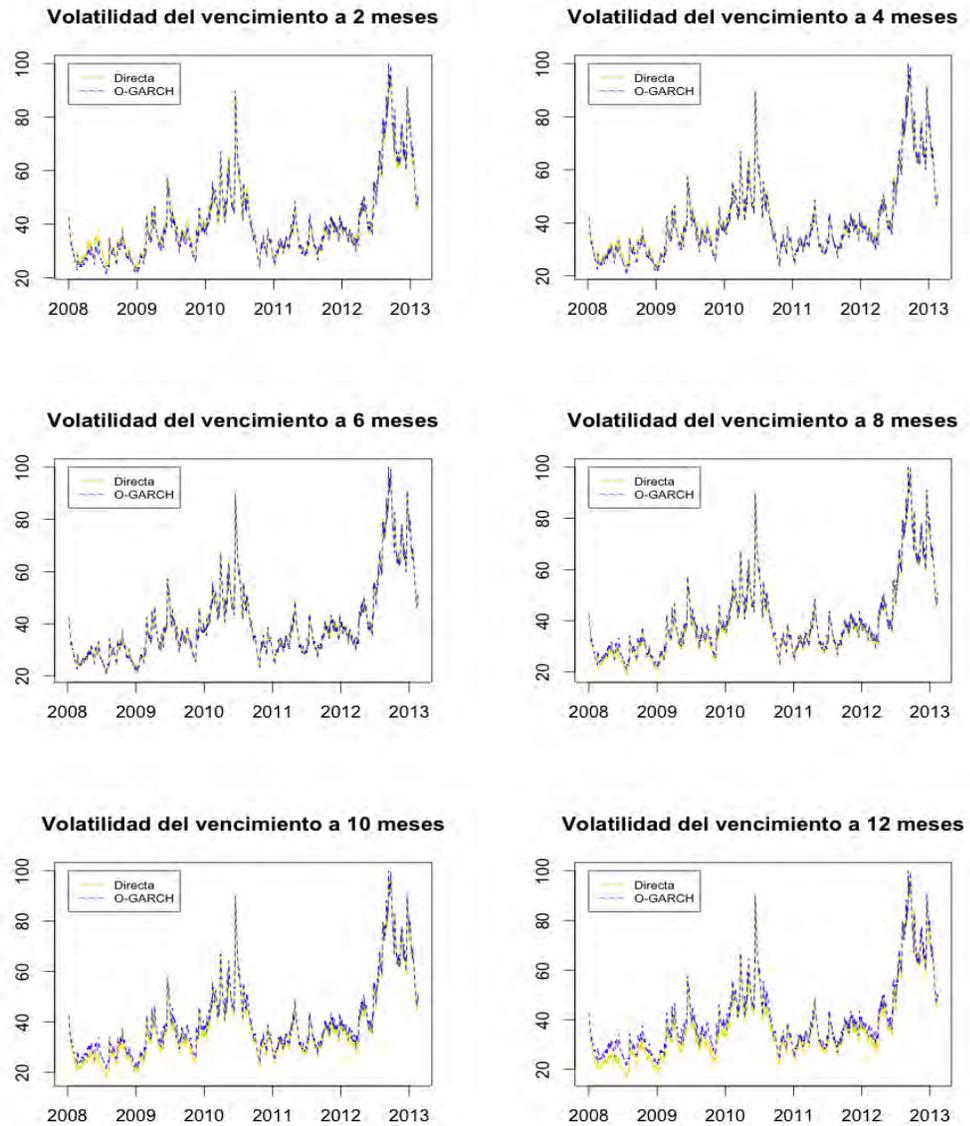
**Tabla 3.10:** Modelos GARCH(1,1) con innovaciones  $t$  de las primeras dos componentes principales, donde  $\nu$  representa los grados de libertad de la distribución  $t$ .

La comparación de las volatilidades obtenidas por el método O-GARCH con aquellas volatilidades que resultan de estimar directamente, estas comparaciones se muestran en la Figura 3.20.

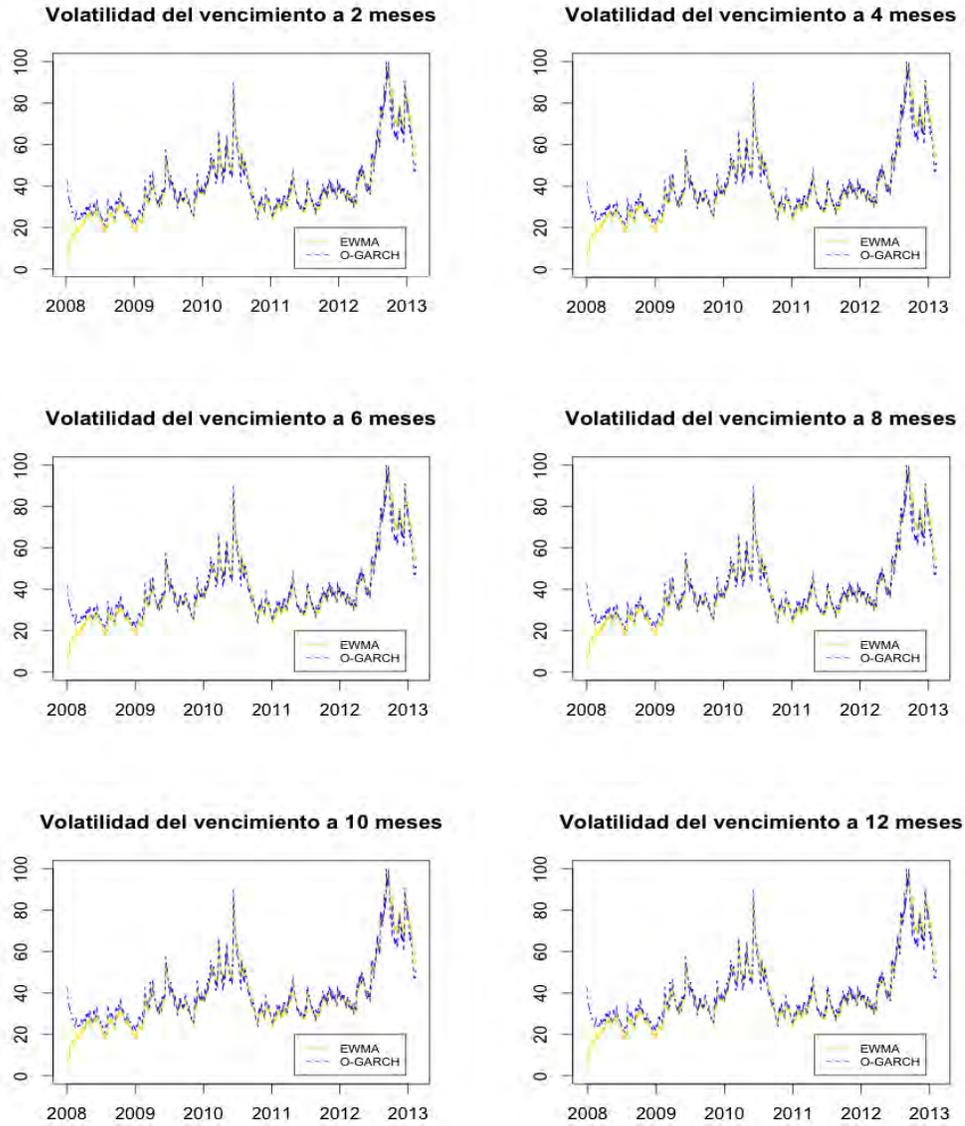
Mientras tanto, en la Figura 3.21, se muestra una comparación del método O-EWMA con el método O-GARCH. Se puede apreciar en ambas gráficas, que las volatilidades son prácticamente idénticas; por lo que es de esperar que nuestra estimación usando el método O-GARCH sea acertada. Tenemos presente que estas comparaciones se hacen para saber si el modelo que estamos utilizando se adecúa realmente a los datos, porque si las volatilidades fueran totalmente distintas, entonces estamos cometiendo un error o utilizando un modelo erróneo..

Por último, revisamos los residuales de las componentes principales para ver si el modelo que ajustamos es correcto. En la Figura 3.22 podemos observar que los residuales se comportan como un ruido blanco estricto, por lo tanto nuestro modelo es correcto.

Era evidente que los *forwards* sobre el *Brent Brand* se comportaran como los *forwards* sobre el WTI, dado que tienen un volumen de negociación muy alto, son muy líquidos y ambos son contratos sobre el crudo del petróleo. Diferente petróleo, pero al fin y al cabo petróleo, ya que éste es un recurso no renovable muy apreciado y requerido por su importancia para el desarrollo económico y social de los países.



**Figura 3.20:** Comparaciones de las volatilidades de los vencimientos 2, 4, 6, 8, 10 y 12 meses de los forwards del petróleo crudo del *Brent Blend*, calculados directamente y usando el método O-GARCH.



**Figura 3.21:** Comparaciones de las volatilidades de los vencimientos 2, 4, 6, 8, 10 y 12 meses de los forwards del petróleo crudo del *Brent Blend*, usando el método O-EWMA y el método O-GARCH.

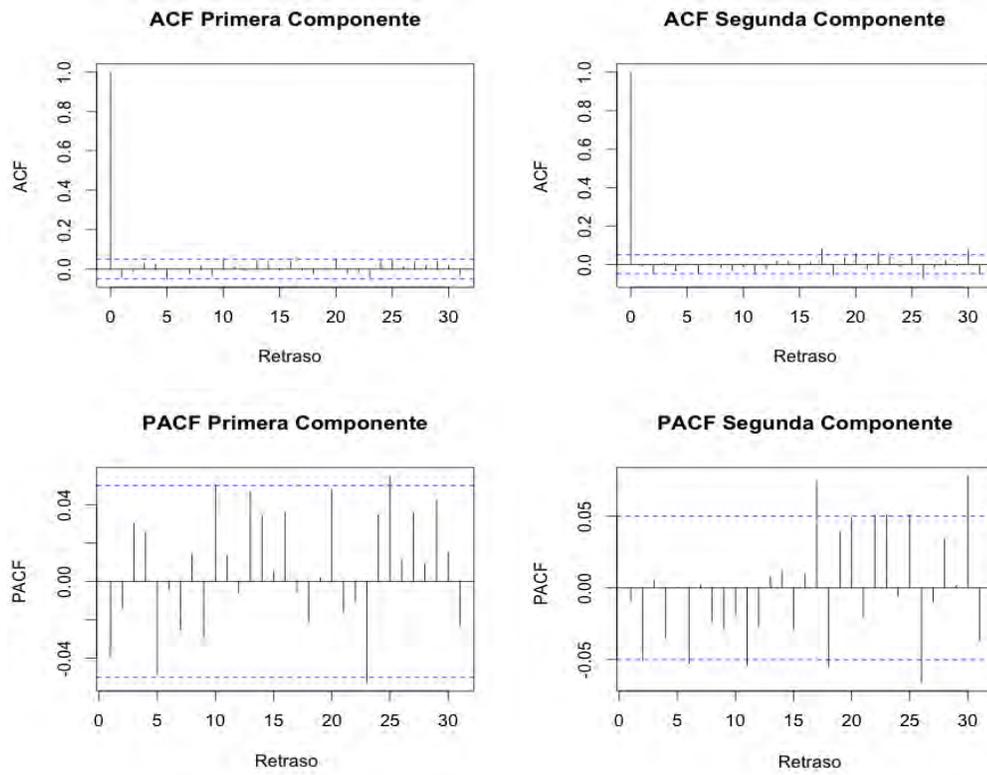


Figura 3.22: ACF y PACF de los residuales de las primeras dos componentes principales.

# Capítulo 4

## Conclusiones

A lo largo de este trabajo, hemos visto algunos de los modelos GARCH Multivariados más importantes, analizamos sus propiedades y también las desventajas que presentan cuando intentamos ajustarlos a datos de la vida real. Dentro de las desventajas más fuertes que presentan estos modelos están el número de parámetros y el ajuste a los datos.

Como vimos en el Capítulo 2, el número de parámetros que necesita el modelo crecerá rápidamente conforme aumente la dimensión, siendo impresionante el caso del modelo VEC que requiere 6105 parámetros cuando estamos en dimensión 10. También está el caso de los modelos CCC-GARCH y DVC que, en la misma dimensión 10, necesitan 75 y 57 parámetros respectivamente, que suena menos descabellado, pero aún tiene muchos parámetros.

Otro problema que presentan, es que al momento de hacer el ajuste de nuestros modelos GARCH Multivariados, porque nos gustaría hacer uso de la función de verosimilitud, como lo hacemos en el caso univariado, pero esto se vuelve cada vez más difícil, por no decir imposible, conforme aumenta la dimensión del modelo. Es por ello que se recurre al ajuste del modelo en etapas, puesto que es nuestra única estrategia viable ante dimensiones muy altas.

Por estas razones que investigamos el método del O-GARCH, ya que éste hace uso del análisis de componentes principales para reducir la dimensión de nuestros datos, pero al mismo tiempo preserva la mayor parte de la varianza del sistema original.

Como observamos en los ejemplos del capítulo anterior, basta con tan sólo 2 varianzas GARCH para calcular, a base de simples transformaciones de la matriz de volatilidades GARCH, las 78 varianzas y covarianzas del sistema de 12 dimensiones. Claro, esto pasará siempre y cuando el análisis de componentes principales se comporte bien, lo cual sucede cuando hay una fuerte correlación en los datos.

Por otra parte, el modelo del O-GARCH permite la estimación por etapas, el cual era otro problema con los otros GARCH Multivariados, y la estimación de este modelo se vuelve bastante sencilla, puesto que estamos utilizando GARCH univariados sobre las componentes principales.

En mi opinión, este modelo ataca el problema multivariado desde otro ángulo, haciéndolo menos complicado que otros modelos y al mismo tiempo, parece ser una extensión natural de los modelos GARCH univariados, puesto que estos son la base para la parametrización de las matrices de varianzas y covarianzas.

A raíz de los trabajos de la Dra. Alexander, surgieron extensiones de este modelo, propuestas por Roy van der Weide y Peter Boswijk, donde en la descomposición espectral de la matriz, en vez de pedir que la matriz  $\Gamma$  sea ortogonal, se pide únicamente que sea no singular, dejando al modelo O-GARCH como una particularización de este método. Este modelo más general lleva el nombre de GO-GARCH (*Generalized O-GARCH*). Este modelo se puede consultar en van der Weide (2002), Boswijk y van der Weide (2006) y Boswijk y van der Weide (2009).

# Apéndice A

## Teoremas

**Definición A.1.** Diremos que una matriz  $A \in \mathcal{M}_{p \times p}$  es ortogonal si cumple que  $A'A = AA' = I_p$ .

**Definición A.2.** Una base ortonormal de  $\mathbb{R}^p$  es aquella que cumple que es base de  $\mathbb{R}^p$  y además todos sus vectores son unitarios y ortogonales entre sí.

**Teorema A.1** (Descomposición Espectral). Toda matriz  $A \in \mathcal{M}_{p \times p}$  simétrica con entradas reales puede ser escrita de la forma

$$A = \Gamma \Lambda \Gamma' = \sum_{j=1}^p \lambda_j \gamma_j \gamma_j',$$

donde

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p),$$

es la matriz diagonal de eigenvalores de  $A$  que, sin pérdida de generalidad, están ordenados tal que  $\lambda_1 \geq \dots \geq \lambda_p$ , y

$$\Gamma = (\gamma_1, \dots, \gamma_p),$$

es una matriz ortogonal formada por los eigenvectores estandarizados de  $A$ .

*Demostración.* Notemos que  $\Gamma$  es una base ortonormal de  $\mathbb{R}^p$ . Por lo que  $\forall i$  se cumple que

$$\Gamma \Lambda \Gamma' \gamma_i = \Gamma \Lambda e_i,$$

porque

$$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_i \\ \vdots \\ \gamma_p \end{pmatrix} \gamma'_i = \begin{pmatrix} \gamma_1 \cdot \gamma'_i \\ \vdots \\ \gamma_i \cdot \gamma'_i \\ \vdots \\ \gamma_p \cdot \gamma'_i \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix},$$

la base de los vectores es ortonormal, *i.e.*  $\gamma_i \cdot \gamma'_j = 0 \forall i \neq j$  y  $\gamma_i \cdot \gamma'_i = e_i \forall i$ , entonces

$$\Gamma \Lambda \Gamma' \gamma_i = \Gamma \Lambda e_i = \Gamma \lambda_i e_i = \lambda_i \Gamma e_i = \lambda_i \gamma_i = A \gamma_i.$$

Por otro lado

$$\left( \sum_{j=1}^p \lambda_j \gamma_j \gamma'_j \right) \gamma_i = \lambda_1 \gamma_1 \gamma'_1 \gamma_i + \dots + \lambda_i \gamma_i \gamma'_i \gamma_i + \dots + \lambda_p \gamma_p \gamma'_p \gamma_i = \lambda_i \gamma_i = A \gamma_i.$$

□

Usando la descomposición espectral, podemos definir de manera sencilla las potencias de las matrices. Sea, por ejemplo,  $A \in \mathcal{M}_{p \times p}$  simétrica, entonces por el Teorema A.1

$$A = \Gamma \Lambda \Gamma',$$

y para alguna  $\alpha \in \mathbb{R}$

$$A^\alpha = \Gamma \Lambda^\alpha \Gamma',$$

donde  $\Lambda^\alpha = \text{diag}(\lambda_1^\alpha, \dots, \lambda_p^\alpha)$ . En particular, ésto nos sirve para calcular la inversa de una matriz, si hacemos que  $\alpha = -1$ .

**Definición A.3.** Una forma cuadrática  $Q(x)$  se forma a partir de una matriz  $A \in \mathcal{M}_{p \times p}$  simétrica y un vector  $x \in \mathbb{R}^p$

$$Q(x) = x' A x = \sum_{i=1}^p \sum_{j=1}^p a_{ij} x_i x_j,$$

y diremos que  $Q(x)$  es positiva definida si

$$Q(x) > 0, \quad \forall x \neq 0,$$

y positiva semidefinida si

$$Q(x) \geq 0, \quad \forall x \neq 0.$$

**Definición A.4.** Una matriz  $A \in \mathcal{M}_{p \times p}$  será positiva definida (semidefinida) si su correspondiente forma cuadrática  $Q(\cdot)$  es positiva definida (semidefinida), y lo denotaremos como  $A > 0$  ( $A \geq 0$ ).

**Teorema A.2.** Si  $A \in \mathcal{M}_{p \times p}$  es simétrica y  $Q(x) = x'Ax$  es su correspondiente forma cuadrática, entonces existe una transformación  $x \mapsto \Gamma'x = y \cdot \exists$

$$x'Ax = \sum_{i=1}^p \lambda_i y_i^2,$$

donde  $\lambda_i$  son los eigenvalores de  $A$ .

*Demostración.* Por el Teorema A.1 tenemos que  $A = \Gamma\Lambda\Gamma'$ , y sea  $y = \Gamma'x$ , entonces  $x'Ax = x'\Gamma\Lambda\Gamma'x = y'\Lambda y = \sum_{i=1}^p \lambda_i y_i^2$ .  $\square$

**Teorema A.3.**  $A > 0$  si y sólo si  $\lambda_i > 0, \forall i = 1, \dots, p$ .

*Demostración.* Por el Teorema A.2

$$0 < \lambda_1 y_1^2 + \dots + \lambda_p y_p^2 = \sum_{i=1}^p \lambda_i y_i^2 = x'Ax \quad \forall x \neq 0.$$

$\square$

Por lo tanto, se puede deducir a partir de los eigenvalores si una forma cuadrática es positiva definida o no.

**Teorema A.4.** Si  $A$  y  $B$  son matrices simétricas y  $B > 0$ , entonces el máximo de  $x'Ax$  bajo la restricción  $x'Bx = 1$  está dado por el eigenvalor más grande de  $B^{-1}A$ . Más general aún,

$$\max_{\{x: x'Bx=1\}} x'Ax = \lambda_1 \geq \dots \geq \lambda_p = \min_{\{x: x'Bx=1\}} x'Ax,$$

donde  $\lambda_1, \dots, \lambda_p$  son los eigenvalores de  $B^{-1}A$ . El vector que maximiza (minimiza)  $x'Ax$  bajo la restricción  $x'Bx = 1$  es el eigenvector de  $B^{-1}A$  que corresponde al eigenvalor más grande (chico) de  $B^{-1}A$ .

*Demostración.* Por definición,  $B^{1/2} = \Gamma_B \Lambda_B^{1/2} \Gamma_B'$ . Sea  $y = B^{1/2}x$ , o bien  $x = B^{-1/2}y$ , entonces

$$\max_{\{x: x'Bx=1\}} x'Ax = \max_{\{x: x'B^{1/2}B^{1/2}x=1\}} x'Ax = \max_{\{y: y'y=1\}} y'B^{-1/2}AB^{-1/2}y.$$

Por el Teorema A.1, la descomposición espectral de  $B^{-1/2}AB^{-1/2} = \Gamma\Lambda\Gamma'$ . Sea  $z = \Gamma'y \Rightarrow z'z = y\Gamma\Gamma'y = y'y$ , por lo que

$$\max_{\{y:y'y=1\}} y'B^{-1/2}AB^{-1/2}y = \max_{\{y:y'y=1\}} y'\Gamma\Lambda\Gamma'y = \max_{\{z:z'z=1\}} z'\Lambda z = \max_{\{z:z'z=1\}} \sum_{i=1}^p \lambda_i z_i^2,$$

siendo la última igualdad por el Teorema A.2. Pero

$$\max_{\{z:z'z=1\}} \sum_{i=1}^p \lambda_i z_i^2 \leq \lambda_1 \max_{\{z:z'z=1\}} \sum_{i=1}^p z_i^2 = \lambda_1,$$

entonces el máximo se obtiene cuando  $z = (1, 0, \dots, 0)'$ , ésto nos da que

$$z = \Gamma'y \Rightarrow \Gamma z = y \Rightarrow \gamma_1 = y \Rightarrow x = B^{-1/2}y = B^{-1/2}\gamma_1,$$

y como  $B^{-1}A$  y  $B^{-1/2}AB^{-1/2}$  tienen los mismos eigenvalores, queda demostrado.  $\square$

# Apéndice B

## Códigos

Basados en los códigos que usa la Dra. Alexander en Alexander (2001), los cuales pueden ser consultados en su página personal <http://www.carolalexander.org/discussionpapers.php>. Los códigos de la Dra. Alexander fueron creados para el lenguaje TSP (*Time Series Processor*), en tanto que los aquí mostrados están hechos para *R Project*.

**Cargamos una base de datos que contiene el número de observaciones de las variables en las columnas, y después calculamos los log-rendimientos.**

```
datos<-read.table("datos.txt",header=TRUE)
nFil<-length(datos[,1])-1
nCol<-length(datos[1,])
logM<-matrix(nrow=nFil,ncol=nCol)
for(i in 1:nCol)
{
  logM[,i]<-diff(log(datos[,i]))
}
```

**Fijamos la constante de suavidad del EWMA a 0.95. Calculamos los productos curzados de los log-rendimientos y después estimamos la volatilidad con el EWMA. Para finalizar guardamos**

las volatilidades y las covarianzas en la matriz **ts**.

```

lambdaEWMA<-0.95
cplogM<-matrix(nrow=nFil,ncol=(nCol*(nCol+1)/2))
for(i in 1:nCol)
{
  cplogM[,i]<-logM[,i]*logM[,i]
}
cont<-0
bool<-FALSE
for(j in 1:nCol)
{
  for(k in 1:nCol)
  {
    for(i in (nCol+1+cont):(nCol*(nCol+1)/2))
    {
      if(j<k && bool==FALSE)
      {
        cplogM[,i]<-logM[,j]*logM[,k]
        cont<-cont+1
        bool<-TRUE
      }
    }
  }
  bool<-FALSE
}
}
sigmaDirEWMA<-matrix(0,nrow=nFil,ncol=(nCol*(nCol+1)/2))
for(j in 1:(nCol*(nCol+1)/2))
{
  for(i in 2:nFil)
  {
    sigmaDirEWMA[i,j]<-(1-lambdaEWMA)*cplogM[i-1,j]
    +lambdaEWMA*sigmaDirEWMA[i-1,j]
  }
}
normal<-apply(sqrt(sigmaDirEWMA*250),2,function(x) max(x))
ts<-matrix(nrow=nFil,ncol=(nCol*(nCol+1)/2))
for(i in 1:nCol)

```

```

{
  ts[,i]<-sqrt(sigmaDirEWMA[,i]*250)/normal[i]*100
}
cont<-0
bool<-FALSE
for(j in 1:nCol)
{
  for(k in 1:nCol)
  {
    for(i in (nCol+1+cont):(nCol*(nCol+1)/2))
    {
      if(j<k && bool==FALSE)
      {
        ts[,i]<-sigmaDirEWMA[,i]
          /sqrt(sigmaDirEWMA[,j]*sigmaDirEWMA[,k])
        cont<-cont+1
        bool<-TRUE
      }
    }
  }
  bool<-FALSE
}
}

```

**Ahora, estimamos la volatilidad con un GARCH(1,1) y guardamos las volatilidades en la matriz tsGARCH.**

```

sigmaDirGARCH<-matrix(0,nrow=nFil,ncol=nCol)
for(j in 1:nCol)
{
  yDir<-garchFit( garch(1,1),logM[,j])
  sigmaDirGARCH[,j]<-yDir@sigma.t2
}
normalGARCH<-apply(sqrt(sigmaDirGARCH*250),2,function(x) max(x))
tsGARCH<-matrix(nrow=nFil,ncol=nCol)
for(i in 1:nCol)
{
  tsGARCH[,i]<-sqrt(sigmaDirGARCH[,i]*250)/normalGARCH[i]*100
}

```

```
}

```

Se hace el análisis de componentes principales y se escoge el número de componentes.

```
PCA<-princomp(logM)
noPCA<-z

```

Se estima la varianza de las componentes principales con el EWMA. Obtenemos la matriz A que tiene los eigenvectores de las componentes principales y después aplicamos el O-EWMA, al finalizar guardamos las varianzas y las covarianzas del método en la matriz tsADA.

```
sigmaEWMA<-matrix(0,nrow=nFil,ncol=noPCA)
for(j in 1:noPCA)
{
  for(i in 2:nFil)
  {
    sigmaEWMA[i,j]<-(1-lambdaEWMA)*(PCA$scores[i - 1, j]^2)
    +lambdaEWMA*sigmaEWMA[i-1,j]
  }
}
A<-PCA$loadings[,1:noPCA]
tsADA<-matrix(nrow=nFil,ncol=(nCol*(nCol+1)/2))
ada<-matrix(0,nrow=nFil,ncol=(nCol*(nCol+1)/2))
for(i in 1:nFil)
{
  ada[i,1:nCol]<-diag(A %* %diag(sigmaEWMA[i,]) %* %t(A))
}
tem<-matrix(nrow=nCol,ncol=nCol)
cont<-0
bool<-FALSE
for(j in 1:nCol)
{
  for(k in 1:nCol)

```

```

{
  for(i in (nCol+1+cont):(nCol*(nCol+1)/2))
  {
    if(j<k && bool==FALSE)
    {
      for(h in 1:nFil)
      {
        tem<-A %* %diag(sigmaEWMA[h,]) %* %t(A)
        ada[h,i]<-tem[j,k]
      }
      cont<-cont+1
      bool<-TRUE
    }
  }
  bool<-FALSE
}
}
normal<-apply(sqrt(ada*250),2,function(x) max(x))
for(i in 1:nCol)
{
  tsADA[,i]<-sqrt(ada[,i]*250)/normal[i]*100
}
cont<-0
bool<-FALSE
for(j in 1:nCol)
{
  for(k in 1:nCol)
  {
    for(i in (nCol+1+cont):(nCol*(nCol+1)/2))
    {
      if(j<k && bool==FALSE)
      {
        tsADA[,i]<-ada[,i]/sqrt(ada[,j]*ada[,k])
        cont<-cont+1
        bool<-TRUE
      }
    }
  }
  bool<-FALSE
}

```

```

    }
  }

```

En esta parte aplicamos el método del O-GARCH, primero ajustamos un GARCH(1,1) a las componentes principales, y con las volatilidades calculamos el resto de las varianzas y covarianzas del sistema, las cuales guardamos en la matriz tsOG.

```

sigmaOG<-matrix(0,nrow=nFil,ncol=noPCA)
for(i in 1:noPCA)
{
  y<-garchFit(~garch(1,1),PCA$scores[,i])
  sigmaOG[,i]<-y@sigma.t2
}
tsOG<-matrix(nrow=nFil,ncol=(nCol*(nCol+1)/2))
adaOG<-matrix(0,nrow=nFil,ncol=(nCol*(nCol+1)/2))
for(i in 1:nFil)
{
  adaOG[i,1:nCol]<-diag(A %* %diag(sigmaOG[i,]) %* %t(A))
}
tem<-matrix(nrow=nCol,ncol=nCol)
cont<-0
bool<-FALSE
for(j in 1:nCol)
{
  for(k in 1:nCol)
  {
    for(i in (nCol+1+cont):(nCol*(nCol+1)/2))
    {
      if(j<k && bool==FALSE)
      {
        for(h in 1:nFil)
        {
          tem<-A %* %diag(sigmaOG[h,]) %* %t(A)
          adaOG[h,i]<-tem[j,k]
        }
      }
      cont<-cont+1
    }
  }
}

```

```

        bool<-TRUE
      }
    }
    bool<-FALSE
  }
}
normal<-apply(sqrt(adaOG*250),2,function(x) max(x))
for(i in 1:nCol)
{
  tsOG[,i]<-sqrt(adaOG[,i]*250)/normal[i]*100
}
cont<-0
bool<-FALSE
for(j in 1:nCol)
{
  for(k in 1:nCol)
  {
    for(i in (nCol+1+cont):(nCol*(nCol+1)/2))
    {
      if(j<k && bool==FALSE)
      {
        tsOG[,i]<-adaOG[,i]/sqrt(adaOG[,j]*adaOG[,k])
        cont<-cont+1
        bool<-TRUE
      }
    }
  }
  bool<-FALSE
}
}

```

**Para finalizar hacemos los plots de nuestras series.**

```

for(i in 1:(nCol*(nCol+1)/2))
{
  plot(ts[,i],type='l',col='gold')
  lines(tsADA[,i],lty=2,col='blue')
  lines(tsOG[,i],lty=2,col='black')
}

```

}

# Bibliografía

- [1] Alexander C., Chibumba A. (1996) *Multivariate Orthogonal Factor GARCH*. University of Sussex Discussion Papers in Mathematics.
- [2] Alexander C. (2000-06) *Orthogonal methods for generating large positive semi-definite covariance matrices*. ISMA Centre Discussion Papers in Finance.
- [3] Alexander C. (2001) *A Primer on the Orthogonal GARCH Model*. ISMA Centre Discussion Papers in Finance.
- [4] Alexander C. (2002) *Principal component models for generating large GARCH covariance matrices*. Economic Notes, 31 (2) pp.337-359.
- [5] Alexander C. (2007) *Moving Average Models for Volatility and Correlation, and Covariances Matrices*. ISMA Centre Discussion Papers in Finance.
- [6] Bauwens L., Laurent S., Rombouts J. V. K. (2003) *Multivariate GARCH Models: A Survey*. Journal of Applied Econometrics.
- [7] Bollerslev T. (1986) *Generalized autoregressive conditional heteroskedasticity*. Journal of Econometrics 31 pp 307-327.
- [8] Boswijk H. P., van der Weide R. (2006) *Wake me up before you GO-GARCH*. UvA-Econometrics Discussion Paper 2006/03.
- [9] Boswijk H. P., van der Weide R. (2009) *Method of Moments Estimation of GO-GARCH Models*. Working paper, University of Amsterdam.
- [10] Brockwell P. J., Davis R. A. (2002) *Introduction to Time Series and Forecasting*. Springer.

- [11] Engle R. F. (1982) *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*. *Econometrica* 50:4 pp 987-1007.
- [12] Flury B., Riedwyl H. (1988) *Multivariate Statistics, A practical Approach*. Cambridge University Press.
- [13] Härdle W., Simar L. (2003) *Applied Multivariate Statistical Analysis*. TECH.
- [14] Johnson R. A., Wichern D. W. (2007) *Applied Multivariate Statistical Analysis*. Pearson.
- [15] Jolliffe I. T. (2002) *Principal Component Analysis*. Springer.
- [16] McNeil A. J., Frey R., Embrechts P. (2005) *Quantitative Risk Management. Concepts, Techniques and Tools*. Princeton Series in Finance.
- [17] Minkah R. (2007) *Forecasting Volatility*. U.U.D.M. Project Report 2007:7.
- [18] Shumway R. H., Stoffer D. S. (2006) *Time Series Analysis and Its Applications With R Examples*. Springer Texts in Statistics.
- [19] van der Weide R. (2002) *GO-GARCH: A multivariate generalized orthogonal GARCH model*. *Journal of Applied Econometrics*, 17, 549-564.
- [20] Zakoïan J. M., Francq C. (2010) *GARCH Models. Structure, Statistical Inference and Financial Applications*. Wiley.